IDENTIFYING HAZARDOUS LOCATIONS BASED ON EXPECTED CRASH FREQUENCY ON YANGON-MANDALAY EXPRESSWAY IN MYANMAR

Cho Thet Mon

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Transportation Engineering

Suranaree University of Technology

Academic Year 2016

การระบุตำแหน่งจุดอันตรายโดยใช้ค่าคาดหวังของความถี่การเกิดอุบัติเหตุบนทางด่วนสายย่างกุ้ง-มัณฑะเลย์ ประเทศเมียนมาร์

โชว์ ตะ มู

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมขนส่ง
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2559

IDENTIFYING HAZARDOUS LOCATIONS BASED ON EXPECTED CRASH FREQUENCY ON YANGON-MANDALAY EXPRESSWAY IN MYANMAR

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee

(Dr. Suthatip Pueboobpaphan)
Chairperson

(Asst. Prof. Dr. Rattaphol Pueboobpaphan)
Member (Thesis Advisor)

(Assoc. Prof. Dr. Vatanavongs Ratanavaraha)
Member

(Asst. Prof. Dr. Vuttichai Chatpattananan)
Member

(Dr. Sajjakaj Jomnonkwao)
Member

(Prof. Dr. Sukit Limpijumnong)
Vice Rector for Academic Affairs and Innovation

(Assoc. Prof. Flt. Lt. Dr. Kontorn Chamniprasart)
Dean of Institute of Engineering

โชว์ ตะ มู : การระบุตำแหน่งจุดอันตรายโดยใช้ค่าคาดหวังของความถี่การเกิดอุบัติเหตุบนทางด่วนสายย่างกุ้ง-มัณฑะเลย์ ประเทศเมียนมาร์ (IDENTIFYING HAZARDOUS LOCATIONS BASED ON EXPECTED CRASH FREQUENCY ON YANGON-MANDALAY EXPRESSWAY IN MYANMAR) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.รัฐพล ภู่บุบผาพันธ์, 81 หน้า

วัตถุประสงค์ของงานวิจัยนี้คือการระบุตำแหน่งของจุดอันตรายบนทางด่วนที่เชื่อมระหว่างเมืองย่างกุ้งกับเมืองมัณฑะเลย์ โดยใช้ค่าคาดหวังจากความถี่ในการชนและการสืบหาความสัมพันธ์ระหว่างความถี่ในการชนกับคุณลักษณะของถนน โดยทางด่วนเส้นดังกล่าวเป็นหนึ่งในเส้นทางที่มีความสำคัญที่สุดของประเทศที่ทำหน้าที่เชื่อมระหว่างเมืองใหญ่สองเมืองและพาดผ่านเมืองเนปิดอว์ ซึ่งเป็นเมืองหลวงของประเทศเมียนมาร์ ซึ่งปัญหาอุบัติเหตุบนทางด่วนกลายเป็นปัญหาใหญ่ที่ส่งผลให้มีผู้บาดเจ็บ พิการ และเสียชีวิตมากมาย รวมทั้งสร้างความเสียหายต่อทรัพย์สินส่วนบุคคลและของสาธารณะ ซึ่งชี้ให้เห็นว่าเป็นปัญหาที่ต้องการการแก้ไขอย่างเร่งด่วนและต้องการให้เป็นไปอย่างมีระบบในเรื่องการปรับปรุงความปลอดภัยบนท้องถนน แต่จนถึงปัจจุบันยังมีการศึกษาวิจัยในเชิงสถิติเกี่ยวกับแนวทางการปรับปรุงความปลอดภัยของทางด่วนเส้นดังกล่าวไม่มากนัก

แบบจำลองการถดถอยทวินามเชิงลบเป็นแบบจำลองที่สามารถทำนายจำนวนครั้งในการเกิดการชนได้จากตัวแปรด้านคุณลักษณะของถนนเช่น ปริมาณจราจรเฉลี่ยใน 1 วัน, ตัวแปรด้านลักษณะเรขาคณิตของถนน จุดที่มีสะพานและแหล่งชุมชนตลอดแนวเส้นทางบริเวณทางด่วน โดยนำข้อมูลของการชนย้อนหลัง 3 ปีนำมาใช้เป็นข้อมูลในการพัฒนาแบบจำลอง จากผลการทำนายของแบบจำลองพบว่า มีความสัมพันธ์ระหว่างความถี่ในการชนกับตัวแปรด้านคุณลักษณะของถนนบางตัวแปรได้แก่ ปริมาณจราจรเฉลี่ยใน 1 วัน, จุดที่มีสะพาน, แหล่งชุมชน, ความลาดชันของถนน, และตัวแปรที่ควบรวมระหว่างโค้งในแนวราบกับความลาดชันของถนน

หลังจากดำเนินการสร้างแบบจำลองคาดการณ์ความถี่ของอุบัติเหตุแล้ว ได้มีการประยุกต์ใช้วิธี Empirical Bayesian (EB) เพื่อประมาณหาค่าคาดหวังของความถี่การเกิดอุบัติเหตุเพื่อให้แบบจำลองมีความแม่นยำมากยิ่งขึ้น และได้ใช้วิธีในการระบุตำแหน่งของจุดอันตรายบนทางด่วนระหว่างเมือง Yangon-Mandalayจำนวน 4 วิธี ได้แก่ Accident Frequency Method, Accident Rate Method, Rate Quality Control Method และ Combined Method ซึ่งในการระบุตำแหน่งของจุดอันตรายจะทำโดยพิจารณาเรียงลำดับของแต่ละช่วงถนนย่อยตามระดับความเสี่ยง โดยช่วงถนนที่มีจำนวนความถี่ของอุบัติเหตุที่สูงกว่าก็จะถูกระบุให้เป็นจุดอันตรายลำดับแรก ๆ

ซึ่งคาดหวังว่าการศึกษานี้จะสามารถนำไปใช้เป็นแนวทางในการจัดลำดับความสำคัญในการปรับปรุงจุดอันตรายบนทางด่วนสาย Yangon-Mandalay ภายใต้เงื่อนไขด้านงบประมาณที่จำกัด

สาขาวิชา <u>วิศวกรรมขนส่ง</u>	ลายมือชื่อนักศึกษา______________
ปีการศึกษา 2559	ลายมือชื่ออาจารย์ที่ปรึกษา ______________
	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม ______________

CHO THET MON : IDENTIFYING HAZARDOUS LOCATIONS BASED ON EXPECTED CRASH FREQUENCY ON YANGON-MANDALAY EXPRESSWAY IN MYANMAR. THESIS ADVISOR : ASST. PROF. RATTAPHOL PUEBOOBPAPHAN, Ph.D., 81 PP.

YANGON - MANDALAY EXPRESSWAY/HAZARDOUS LOCATIONS/ EXPECTED CRASH FREQUENCY/EMPIRICALBAYESIAN/NEGATIVE BINOMIAL REGRESSION

The purpose of this research is to identify hazardous locations base on expected crash frequency and to investigate relationships between crash frequencies and road characteristics on Yangon-Mandalay Expressway in Myanmar. This expressway is the most important one in the country since it connects between two big cities, Yangon and Mandalay, and passes through the capital city of Myanmar, which is named Naypyitaw. Traffic crashes on that Expressway become a worse problem resulting in many deaths, injuries, disabilities and damage to both private and public properties. These losses point out that there is an urgent need for a systematic approach to improve road safety. Up to present, there have not been much statistical researches conducted on the topic of road safety improvement for this expressway.

Negative binomial regression model was performed to predict the numbers of crash based on road characteristics variables. These variables include Average Daily Traffic, road geometric variables, presence of bridge and presence of village settlement along the expressway. The last three years traffic crash data were used to develop the crash prediction model. According to the predictive modeling result, it was found that there are relationships between crash frequencies and some road

characteristics variables; Average Daily Traffic, presence of bridge, presence of village settlement, percent downgrade and combination of horizontal curve and slope.

After performing crash prediction model, Empirical Bayesian (EB) estimation method was then applied to improve the precision of predicted crash frequency. Finally, four methods, namely Accident Frequency Method, Accident Rate Method, Rate Quality Control Method, and Combined Method, were carried out to identify hazardous locations on Yangon-Mandalay expressway. In hazardous identification section, all road segments were ranked according to their degree of risk and the segment with highest number of crashes was determined more dangerous and ranked as first hazardous location. It is expected that this study will be useful for prioritizing safety treatment for hazardous locations on Yangon-Mandalay expressway under limited budget condition.

School of Transportation Engineering Student's Signature ________________

Academic Year 2016 Advisor's Signature ________________

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my advisor, Asst.Prof. Dr. Rattaphol Pueboobpaphan and Co-advisor, Assoc.Prof.Dr.Vatanavongs Ratanavaraha, for their guidance, motivation and friendly advice during the thesis work. You all definitely provided me with the tools that I needed to choose the right direction and successfully complete my thesis.

Besides my advisors, I would like to thank my thesis committee, Dr. Suthatip Pueboobpaphan, Asst. Prof. Dr. Vuttichai Chatpattananan, Dr. Sajjakaj Jomnonkwao for their encouragement, insightful comments, and hard questions.

My sincere thanks also go to Thailand International Development Cooperation Agency (TICA), for offering me the scholarship to continue my post graduate study. Special thanks are extended to the personnel from the Ministry of Construction (Myanmar) for their help in collection of data.

Finally, I must express my heartfelt gratitude to my family members for providing me with constant support and encouragement throughout the years of study.

This accomplishment would not have been possible without them. Thank you.

Cho Thet Mon

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

Table **Page**

LIST OF TABLES (Continued)

Table **Page**

LIST OF FIGURES

Figures **Page**

SYMBOLS AND ABBREAVTIONS

α	=	Alpha
β	=	Beta
λ	=	Lambda
AADT	=	Annual Average Daily Traffic
AASHTO	=	American Association of State Highway and Transportation Officials
ADT	=	Average Daily Traffic
APM	=	Accident Prediction Model
CPM	=	Crash Prediction Model
EB	=	Empirical Bayes
GIS	=	Geographic Information System
GLM	=	Generalized Linear Model
HI	=	Hazardous Index
SPF	=	Safety Performance Function
SPSS	=	Statistical Package for the Social Sciences

CHAPTER I

INTRODUCTION

1.1 Rationale and Background

Road traffic accidents cause injuries, death and losses of properties; it is one of the problems faced by modern societies of the world today. According to the report of World Health Organization (2013), about 1.24 million people globally die each year as complication of road traffic crashes. There are nearly 3400 deaths a day. Over 90% of the world’s road traffic fatalities occur in low- and middle-income countries, even though these countries have only about half of the world’s vehicles. Without taking action, annual road traffic deaths are likely to increase up to 1.9 million by 2030 and might even become the seventh leading cause of death. Therefore, road traffic crashes are prone to be a major socio-economic problem of the world.

In Myanmar, one of the developing countries in Southeast Asia, a new expressway was constructed in 2005. It passes through several cities including Naypyitaw, the capital of Myanmar, as well as linking Yangon to the south and Mandalay to the north. This expressway is one of the road infrastructure development projects undertaken by the former military government in the country and about 40 miles shorter than the existing old highway. It has two traffic lanes for each direction, 366 miles length, with (30 ft) raised median between the direction of traffic flow. Motorcycle and pedestrian were prohibited to access on the expressway. A few funds were invested in safety measure and engineers were instructed by the government to complete the project in a short period. As a result, the project led to be a rush job. Although it was expected to be the most convenient and safest expressway in Myanmar, many accidents have been occurring on that way. Annual number of traffic crash on that expressway has been increasing with the rapid growth of vehicle ownership. Location of Yangon-Mandalay Expressway is shown in figure1.1.

1.1.1 Specifications of Yangon-Mandalay Express way

- A total length of expressway is 366 mile (589km).
- Road surface comprises two layers of concrete, 12 inches thick surface layer and 6 inches thick lean concrete layer.
- The expressway can withstand vehicle load up to 80 tons and vehicle speed limit is 100 km/hr.
- It has four lane of traffic with 12.5 ft of each and 30 ft wide raised median.
- There are a total of 905 box culverts, 462 bridges and 116 underpasses on the expressway.

1.2 Problem Statement

In all traffic accident reports of Yangon-Mandalay Expressway, human were mostly criticized (due to over speeding) without thoroughly analyzing other factors such as road surface defect, defective geometric design, insufficient number of lanes and lane width , structural deformation of pavement and other road safety management system. Although the major causes of death and level of injury related to traffic crashes on Yangon-Mandalay expressway is due to over speeding and non-wearing of seat-belts, other accident contribution factors such as vehicle error, road environment and road characteristic are also investigable. In all developed and some developing countries, they generally consider traffic accident based on human characteristics, road characteristics, vehicle characteristics and environmental impact. In several previous researches, it has been observed that road characteristics have significant effects on traffic accident. All accidents in Yangon-Mandalay Expressway are listed as due to “driver error”, “over speeding”, “tire bursting” is not a good conclusion that other road characteristics are not involved and no further considerations are required. Figure 1.2 and 1.3 describe the current issue that indicating causes of traffic crash on Yangon-Mandalay Expressway based on the accident report (from 2013 to 2014) recorded by policemen. Figure 1.4 and 1.5 represents the types of accidents happened in 2013 and 2014 respectively.

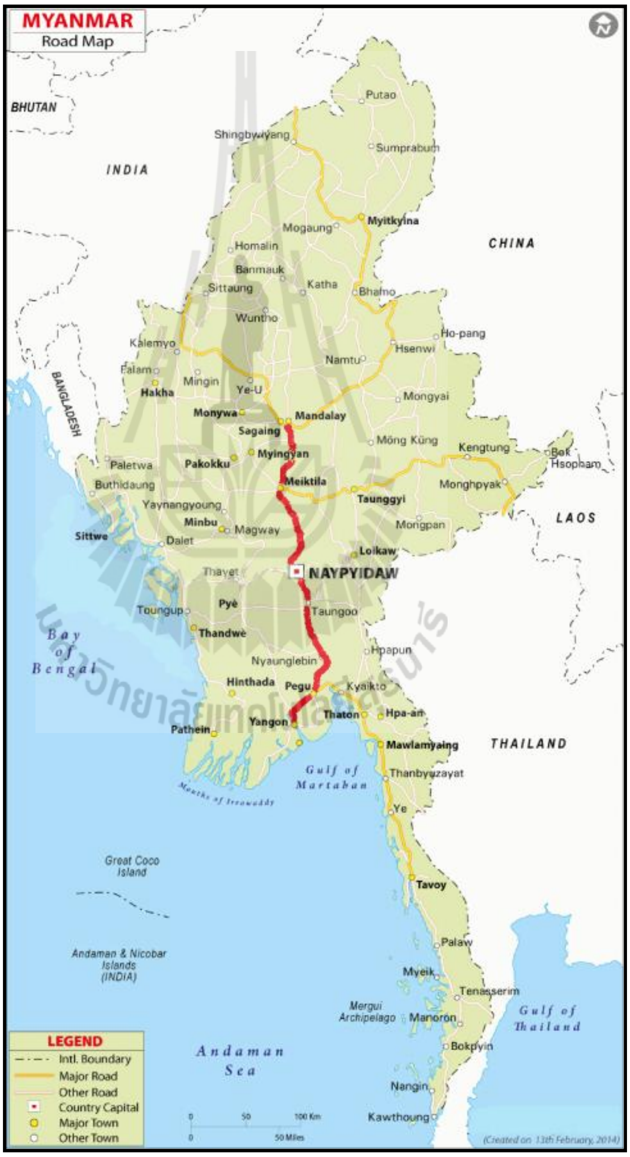
There has not been much statistical research on the topic of traffic crashes due to road characteristics on Yangon-Mandalay expressway as well as identifying hazardous locations. This might have been a result of inadequate information available on road characteristics, road accidents and its impact on human’s lives and properties in the country. For that reason, there is an urgent need to determine traffic crashes related to road characteristics on that expressway. In addition, there is in need of establishing affective procedure so as to identify hazardous highway locations based on accident statistics to support highway safety improvement program.

1.3 Thesis Objective

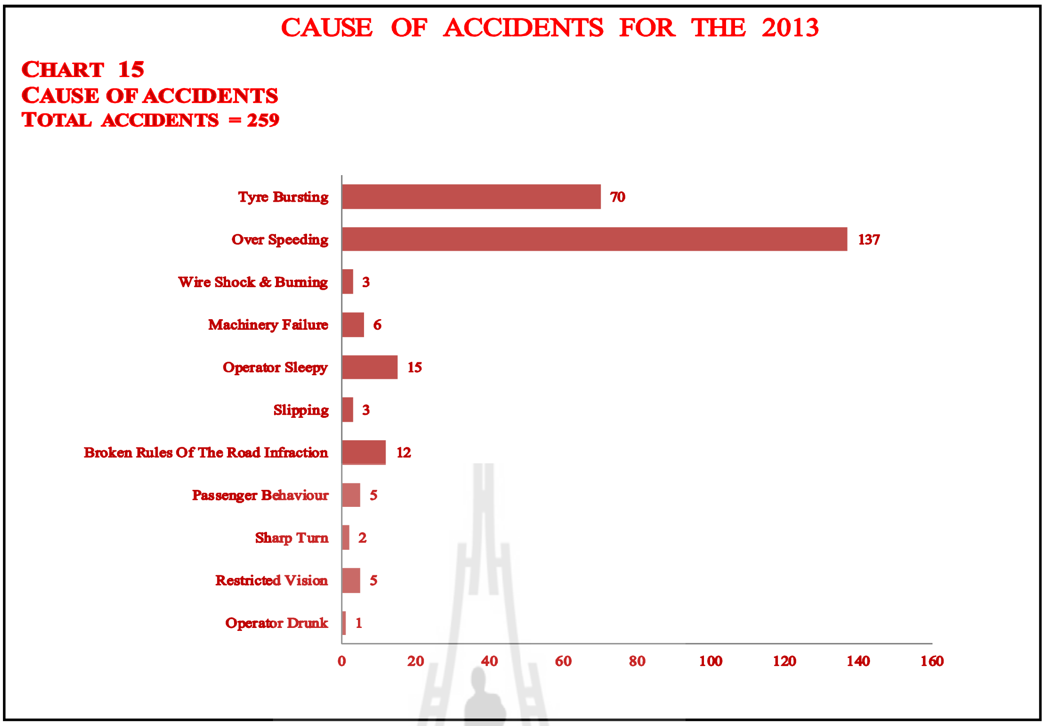
The major objectives of the study are:

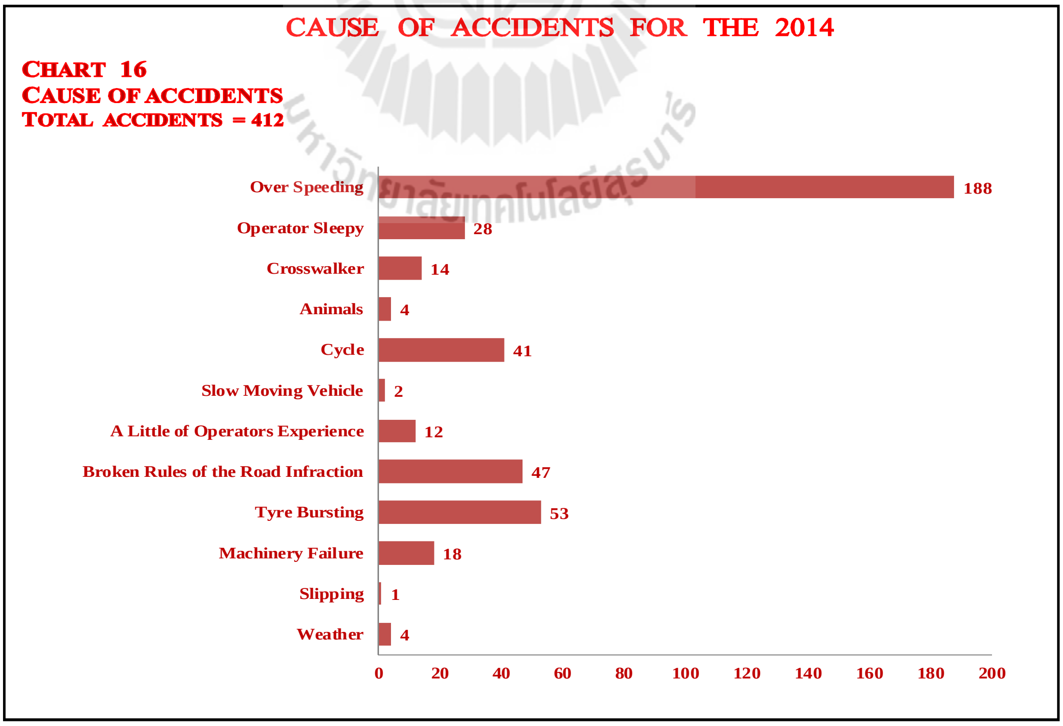
1 To identify hazardous highway locations and

2 To investigate relationships between road characteristics and traffic crash on Yangon-Mandalay Expressway.



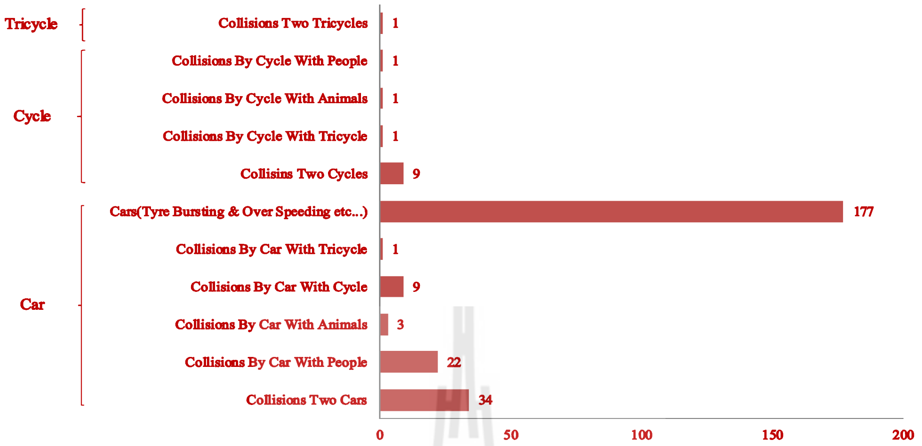
MYANMAR
Road Map
BHUTAN
INDIA
CHINA
BANGLADESH
LAOS
THAILAND
Putao
Shingbwiyang
Sumprabum
Myitkyina
Mogaung
Homalin
Banmauk
Katha
Bhamo
Sittaung
Wuntho
Kalemyo
Ho-pang
Hsenwi
Falam
Namtu
Hakha
Mingin
Ye-U
Mongyai
Monywa
Mandalay
Sagaing
Möng Küng
Kengtung
Myingyan
Paletwa
Pakokku
Meiktila
Hsopham
Buthidaung
Monghpyak
Yaynangyoung
Taunggyi
Minbu
Magway
Mongpan
Sittwe
Dalet
Loikaw
Thayet
NAYPYIDAW
Toungup
Pyè
Taungoo
Thandwe
Bay of Bengal
Hpapun
Nyaunglebin
Hinthada
Pegu
Kyaikto
Thaton
Hpa-an
Yangon
Pathein
Mawlamyaing
Gulf of Martaban
Thanbyuzayat
Mouths of Irrawaddy
Ye
Great Coco Island
Tavoy
Palaw
Andaman & Nicobar Islands (INDIA)
Myeik
Tenasserim
Mergui Archipelago
Manoron
Gulf of Thailand
Bokpyin
Andaman Sea
Nangin
Kawthoung
LEGEND
Intl. Boundary
Major Road
Other Road
Country Capital
Major Town
Other Town
0 50 100 Km
0 50 Miles
(Created on 13th February, 2014)


CAUSE OF ACCIDENTS FOR THE 2013
CHART 15
CAUSE OF ACCIDENTS
TOTAL ACCIDENTS = 259
Tyre Bursting 70
Over Speeding 137
Wire Shock & Burning 3
Machinery Failure 6
Operator Sleepy 15
Slipping 3
Broken Rules Of The Road Infraction 12
Passenger Behaviour 5
Sharp Turn 2
Restricted Vision 5
Operator Drunk 1
0 20 40 60 80 100 120 140 160


CAUSE OF ACCIDENTS FOR THE 2014
CHART 16
CAUSE OF ACCIDENTS
TOTAL ACCIDENTS = 412
Over Speeding 188
Operator Sleepy 28
Crosswalker 14
Animals 4
Cycle 41
Slow Moving Vehicle 2
A Little of Operators Experience 12
Broken Rules of the Road Infraction 47
Tyre Bursting 53
Machinery Failure 18
Slipping 1
Weather 4
0 20 40 60 80 100 120 140 160 180 200

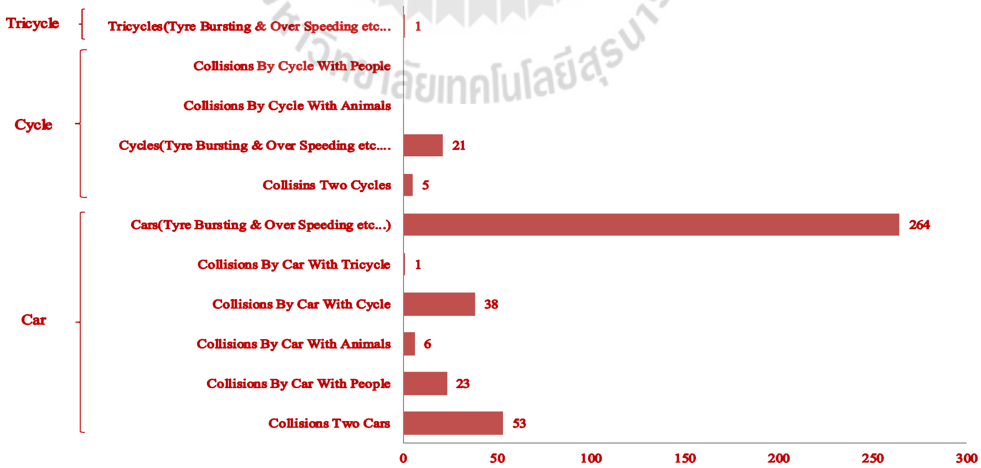
TYPE OF ACCIDENTS FOR THE 2013

CHART 14
TYPE OF ACCIDENTS
TOTAL ACCIDENTS = 259



TYPE OF ACCIDENTS FOR THE 2014

CHART 15
TYPE OF ACCIDENTS
TOTAL ACCIDENTS = 412



1.4 Thesis Question

Is there any relationship between traffic crash occurrences and road characteristics on Yangon-Mandalay Expressway?

1.5 Scope and limitation of the Study

Traffic accidents generally take place due to a number of factors as follow;

- Human related factors
- Road related factors
- Vehicle related factors
- Environmental related factors and
- Their interrelationship.

Among the above factors, the study mainly focused only on the road related factors on Yangon-Mandalay Expressway due to the limitation of data available.

1.6 Expected Contribution

The research finding will contribute to understanding the relationships between road characteristics and traffic crashes and will contribute to provide proper treatments in hazardous locations on Yangon-Mandalay Expressway under limited budget condition. Furthermore, it will contribute to increase public awareness that which parts of the road are dangerous and not dangerous along the expressway.

1.7 Thesis Structure

This thesis consists of six chapters. Chapter One contains a background, a statement of the problem, thesis objectives, thesis question, scope and limitation, excepted contribution and structure. Chapter Two contains literature review of previous work on the study. Chapter Three describes data collection, the methodology and procedures adopted for the study. Analysis and result interpretation are provided in Chapter Four. Chapter five concludes the study followed by recommendation.

CHHAPTER II

LITERATURE REVIEW

2.1 Introduction

There have been analyzing identifying hazardous highway locations based on expected numbers of crash in previous research. The nature of these research papers vary with different types of roads, different numbers of variables used in the crash prediction model and different methods of analyses. On the base of literature review for the study, the followings are the review of related literatures.

2.2 Accident Causative Factors Overview

2.2.1 Traffic Volume

Many researchers have considered Average Daily Traffic (ADT) in an accident prediction model as the most effective parameter and it has positive relationship with traffic accidents. In the study of Saffarzadeh and Pooryari (2005), the result showed that if traffic volume increases by 3 percent, accident probabilities will increase by 100 percent. Sloth and Vingaard (2013) have performed a statistical analysis of accident density as a function of Annual Average Daily Traffic (AADT) and other variables related to road characteristics. Their conclusion is that ADT has positive regression parameter and statistically significant, means that accident frequency will increase if traffic volume increases. Similar result, higher number of crashes with higher traffic volume, can be explored out in the paper of Mohammadi and Samaranayake (2014).

2.2.2 Geometric Characteristics of Roadway

There are many road geometric factors that influence on traffic crashes. A road horizontal curve element is the one among these effective factors. Lesser degree of horizontal curves (i.e. sharp curve) is the possible geometric features that increase the frequency of accidents (Mohamed, 1999). Aram (2010) claimed the importance of horizontal curve of higher crash rates than straight sections of similar length and traffic composition. The increase in crash rate becomes particularly significant at radii below 200m, he also reviewed that horizontal curves are more dangerous when combined with gradients and surface with low coefficients of friction. Hanno (2004) stated that combined crest curves are found to be more collision prone than combined sag curves and avoiding vertical/horizontal curves combination enhances safety. In the previous paper, Bauer and Harwood (2013), their accident prediction model indicated that crash frequency increases with decreasing horizontal curve radius and increase with increasing percent grade and grade difference. And they also stated that short horizontal curve at sharp crest and sag vertical curves are associated with higher crash frequencies. Kontaratos at el., (1994) determined the effect of grade in defining the minimum horizontal curve radius. They found out a strong relationship between the radius of horizontal curve and grade; there should have larger minimum curve radius on longitudinal grade at higher vehicle speed. Amgalan and Jigjjav (2013) claimed that vertical (crest and sag) curves influence on traffic safety negatively because of inadequate visibility.

2.2.3 Road environment factors

Traffic accidents have been found to be associated with other road environment factors. The variable, presence of village settlement, represents the level of pedestrian interference with vehicles on the highway sections and the section which had a village settlement were found to increase injury crashes by 60.3% compared with sections with no settlements (Ackaah and Salifu, 2011). Mohammed et al., (2015) also explained that numbers of towns / villages on the routes seem to play an important role in traffic safety. There have very few recent published literature sources on traffic accident and presence of bridge along the expressway. Ogden (1989) reported about bridge crash prediction model and reviewed significant factors associated with bridge crashes.

2.3 Crash Prediction Model

An accident prediction model (APM) is a mathematical formula indicating the relation between the safety level of existing roads (i.e. crashes, victims, injure fatalities etc) and variables explain this level (road length, width, traffic volume etc). Traffic volume (vehicle per day) and road length are the most important explanatory variable in an APM. The parameter of the model, however, can vary between types of roads and countries. The reason is that the road characteristics, road user behavior, vehicle type and environment of the road can differ from place to place. Therefore APM should be developed per country and road type in order to determine the safety level of particular road (Eenink et al., 2008).

In statistics, the Generalized Linear Model (GLM) is a flexible generalization of ordinary linear regression. GLM allows response variables that have error distribution other than a normal distribution. Currently, (GLM) is the popular approach for the development of Crash Prediction Models. There are a number of crash prediction models in several researches to analyze safety performance of the road (Greibe, 2003), (Ackaah and Salifu, 2011), (Kumar et al., 2013) and (Zou et al., 2015). Among them, Poisson regression is the most widely used in modeling count data since accident data are in the form of count and discrete property. However, the use of Poisson regression has been restricted by a limitation; Poisson regression has a strong assumption of mean equal to variance (Miaou, 1993). Negative Binomial (NB) Model is the extension of Poisson model and an alternative approach to modeling over-dispersion (variance > mean) in count data. it has been remained as the most commonly used statistical tool among several statistical models that have been proposed for modeling crash data (Kibar et al.,2013), (Zou et al., 2015) and (Naznin et al.,2016). Oppong (2012) also argued that NB model is best fit for the over-dispersion data. Although NB is more general than the Poisson model, it cannot perform well when the data is under-dispersed (variance < mean) or characterized by low sample mean and small sample size. Conway Maxwell Poisson Model is one of the integrated Poisson model. The main advantage of this distribution over other model is its ability to handle both the over-dispersion and under-dispersion data. However, it is also noted that the model does not perform well when the mean of the sample is low or the sample size is very small (Lord et al., 2008).

2.4 Identifying Hazardous road location

Identifying hazardous location is an essential first step of road safety management program. It can be carried out in different ways. Although there is no universal recognized definition of a hazardous location (or) black spot (or) hot spot, these locations can generally be determined as locations of higher expected number of accident than other similar locations due to local risk factors.

Several alternative methods have been developed to identify black spot in previous research papers. Zegeer (1982), Gharaybeh (1991) and Brown (1992) applied several techniques to identify hazardous locations. These techniques were accident frequency method, accident rate method, accident severity method, rate quality control method and others. Applying these methods separately can limit overall accuracy of the results. Therefore, a collection or combination method has been developed for the better result of accuracy (Utainarumol and Stammer, 1999). Ratanavaraha and Amprayn (2003) investigated the causative highway accident factors for the first stage expressway system in Thailand. Five methods (accident frequency method, accident rate method, quality control method, accident severity method and combined method) of identifying hazardous locations were performed in their research.

Up to present, GIS (Geographic Information System) is a very popular information management technique and provides several capabilities that can identify hazardous highway locations graphically. Identifying hazardous road location with the use of GIS application can be found in papers of (Utainarumol and Stammer, 1999), (Ahsan and Newaz, 2011) and (Rachman and Newaz, 2013).

According to improvement of hazardous roadway segment identification, the sliding-window is a method for selecting an accident with frequent intervals while overlapping a certain length. Although this method is a progressive method for selecting hazardous zone, it does not have even optimal value and an alternative for the length and placement method (Lee et al., 2013). Elvik (2007) also argued that the use of sliding window approach artificially inflates variation in accident counts.

2.4.1 State-of-the-art Approach for Identifying Hazardous road location

Apart from the above study, Elvik (2007) also surveyed about operational definition of hazardous locations in some European countries and finally proposed characteristics of the state-of –the –art approach for identification of hazardous road locations. These characteristics are as following;

1. Hazardous road locations should be identified from a population of sites whose members can be enumerated.

2. Hazardous road locations should not be identified by applying a sliding window approach.

3. Hazardous road locations should be identified in terms of the site-specific expected number of accidents. The empirical Bayes (EB) method is best suited for obtaining estimates of the expected number of accidents for each site. Use of the EB-method should rely on an accident prediction model.

4. Hazardous road locations should be identified as the upper percentiles in a distribution of the EB-estimates of site-specific safety in a population of sites. They should not be identified in terms of an excessive number of accidents above a certain “normal” level.

5. A suitable period of data for developing an accident prediction model and identifying hazardous road locations is 3–5 years.

6. Accident severity can be considered when identifying hazardous road locations, provided EB-estimates of the site-specific expected number of

accidents by severity can be obtained. The recorded distribution of accidents by severity should not be used.

7. Specific types of accidents can be considered when identifying hazardous road locations, provided EB-estimates of the site-specific expected number of accidents of each type can be obtained.

According to his statement, it can be summarized that hazardous locations should be identified in terms of the site-specific expected number of crash. In addition, the empirical Bayes (EB) method is best suited for obtaining estimates of the expected number of accidents for each site and the use of the EB-method should rely on a crash prediction model.

2.4.2 Empirical Bay (EB) Estimation Method

The empirical Bayes (EB) approach is currently the standard method for estimating traffic safety (Shin, 2008). This is recommended as the state-of -the -art -approach to identify hazardous locations and can increase the precision of estimation and corrects for the regression-to-mean bias (Elvik, 2007). The theory is based on the acceptance that accident counts are not the only key to the safety of an entity; another key is in what is known about the safety of similar entities. A rational estimate must be a mixture of the two keys. The theoretical framework for combining the information contained in accident counts with the information contained in knowing the safety of similar entities is the EB method (Hauer et al., 2002). The EB method use a weight factor, which is a function of over dispersion parameter from a crash prediction model, to combine observed crash frequency and estimated crash frequency from a predicted model into a weighted average (AASHTO, 2010). Valentova and Ambros (2013) performed the comparison between the traditional and

EB approach for identification of hazardous road locations and claimed that the prediction model with EB adjustment offered the result which was more stable in time compared to the traditional approach.

2.5 Summary of the Literature Review

The main purpose of the study was to identify hazardous road locations and to investigate relationships between crash frequencies and road characteristics on Yangon-Mandalay expressway. To deal with this purpose, it is necessary to establish a proper methodology. In proceeding discussion of the literature review, the author has already reviewed about identifying hazardous road locations and state - of - the - art approach of it. Generally, identifying hazardous locations can be carried out in different ways. Of them, GIS application, sliding window method and five methods which we called (Accident Frequency Method, Accident Rate Method, Rate Quality Control Method, Accident Severity Method and Combined Method) have been reviewed in the literature of this study. Other methods might be left to be reviewed by the author in the literature review section. To identify hazardous locations, some criteria of the state-of -the-art approach were considered in the study rather than using traditional method. To overcome this point, an accident prediction modal and Empirical Bay (EB) estimation method were carried out in order to predict expected numbers of accident. And then four methods (Accident Frequency Method, Accident Rate Method, Rate Quality Control Method and Combined Method) were performed for identification of hazardous locations.

CHAPTER III

METHODOLOGY

3.1 Data Collection and Analysis

The data were specified into three groups; crash data, road data and road geometric data. Crash data consists of several types of collision based on the 2013 - 2015 period. The road data comprises information on AADT, surface roughness and surface type of the road, lane and median width, terrain type of the road, location of bridges, locations of village settlement and locations of guardrails. And the road geometric data contains horizontal and vertical profile of the expressway. Secondary types of data were acquired from the Ministry of Construction of the Myanmar government. The data were then extracted and processed by using Microsoft Excel application to perform the crash prediction model.

3.1.1 Crash data

In developed and some developing countries, traffic crash data are collected by responsible organization. Their data collecting systems are of well-specified standards and detailed procedures. In Myanmar, traffic crash report forms are usually filled by policemen. Therefore accident report forms stated by policemen were the only source for obtaining crash data. They cannot thoroughly understand about the trends and patterns of traffic crash, road safety and road features. Consequently, traffic crash data are lack of shortage such as:

- The form does not hold adequate information which is needed for traffic safety analysis.

- Police officers without engineering sense usually adopt the crash report filling process, therefore sometimes; they do not fill report form perfectly.
- Most of the accident causes were assumed to have been speeding.

Crash data received for analysis of study were in the form of paper document. In these documents, crash were identified based on the information; such as (a) date, time and location of the crash, direction of travel (b) driver's name, vehicle type and license plate number, (c) Assumption of crash occurrence, total number of vehicle involved in the crash and numbers of fatal and injured person involved in the crash . As a result of the reason that has been mentioned, collecting and managing crash data for the study was a difficult process.

Motorcycle and pedestrian are not allowed to access on the Yangon-Mandalay expressway, but some road users do not follow the vehicle access regulations. Hence, the crash data contain numerous sources of accidents; include (motorcycle collisions, single vehicle collisions, collisions involving two vehicles, collisions with animals, collisions involving vehicle and pedestrians and collisions due to vehicle breakdown and mechanical failure). Motor vehicle collisions due to vehicle breakdown and mechanical failure were extracted from the recorded crash data and were not considered in analysis process.

3.1.2 Road data

All of the road data were received in hard copy (paper document) format and the first step was to consider which variables were the most suitable for the study. To obtain annual average daily traffic for individual road segments, the following types of data were collected from the Ministry of Construction.

1. Seven days traffic survey carried out at the specific locations of the expressway.

2. Yearly traffic volume (for 2013, 2014, and 2015) that has been recorded by toll stations located along the expressway.

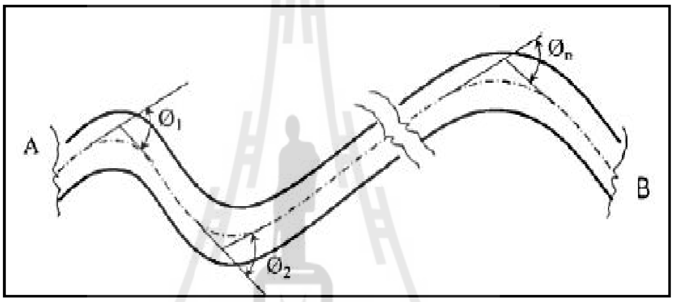
A comparison between the above two types of traffic volume data had been made to find out the reliable value of AADT for the study.

There are a total of 905 box culverts, 462 bridges and 116 underpasses along the Yangon-Mandalay expressway. To figure out relationship between presence of bridge and crash frequency, all of existing box culverts and underpasses were considered as bridge. The variable used in the crash prediction model of the study, presence of bridge, was categorized into three levels and these levels were (1) no bridge presence, (2) one bridge presence and (3) two and more bridges presence.

For determining presence of village settlement along the expressway as a predictor variable in the model; all villages located within 0.6 mile from the right of expressway were considered.

3.1.3 Road geometric data

Road geometric data consists of horizontal and vertical profile of the expressway and these are soft copy (scanned document) format. The total number of horizontal curves on the expressway was (293), their radius ranging from (170) to (5000) meter. (59) percent of these curves had a radius between (170) and (600) meter, and (41) percent of the radii values were between (600) and (5000) meter respectively. The curves having radii value less than (600 meter) were considered as sharp horizontal curves for the study. The horizontal curvature of each road segment for


A
Ø1
Ø2
Øn
B

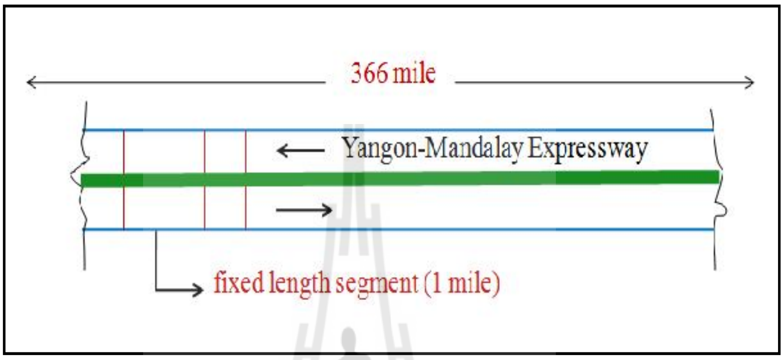

366 mile
Yangon-Mandalay Expressway
fixed length segment (1 mile)

Table 3.1 Variable Selection

No	Type of variable	Variable description
1	Response variable	Number of accident in three years per road segment
2	Predictor variables	Annual Average Daily Traffic per road segment
3		Presence of sharp horizontal curve per road segment
4		Average horizontal curvature per road segment
5		Percent upgrade per road segment
6		Percent downgrade per road segment
7		Combination of horizontal curve and slope per road segment
8		Presence of bridge per road segment
9		Presence of village settlement per road segment

3.4 Variable Explanation on Safety Aspect

This section explains the nature of predictor variables used in the crash prediction model and their influences on the road safety.

3.4.1 Annual Average Daily Traffic and Average Daily Traffic

Annual average daily traffic, abbreviated AADT, is the total volume of vehicle traffic of a highway or road for a year divided by 365 days. AADT is a useful and simple measurement of how busy the road is. One of the most important uses of AADT is for determining funding for the maintenance and improvement of Highways.

Average Daily Traffic ,abbreviated as (ADT),is the average 24 hours traffic volume ,being the total volume during a stated period divided by the number of days in that period. The relationship between traffic volume and the number of accident has been evaluated in a large number of previous studies. It is also found that traffic volume and number of accidents is directly proportional to each other. Traffic volumes not only have an effect on road safety, it is also necessary for the determination of accident rates.

3.4.2 Horizontal Curve and it’s Radius

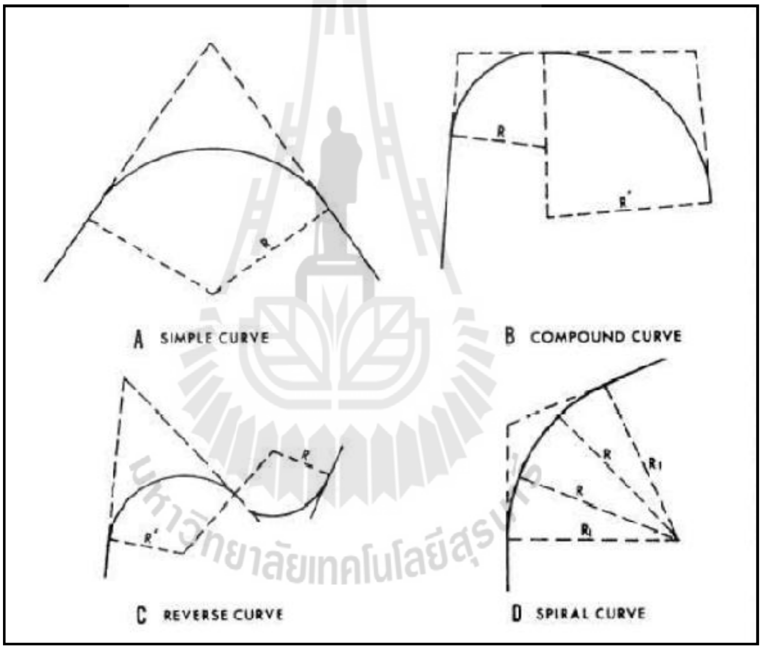
Horizontal curves are the one of important elements in geometric design of highway. A horizontal curve provides a transition between two tangent strips of roadway, allowing a vehicle to negotiate a turn at a gradual rate rather than a sharp cut. The design of the curve is dependent on the intended design speed for the roadway, as well as other factors including drainage and friction. These curves are semicircles as to provide the driver with a constant turning rate with radii determined by the laws of physics surrounding centripetal force.

3.4.2.1 Types of Horizontal Curves

There are four types of horizontal curves and they can be described as follows:

1. SIMPLE: The simple curve is an arc of a circle. The radius of the circle determines the sharpness or flatness of the curve.

2. COMPOUND: Frequently, the terrain will require the use of the compound curve. This curve normally consists of two simple curves joined together and curved in the same direction.



A SIMPLE CURVE
B COMPOUND CURVE
C REVERSE CURVE
D SPIRAL CURVE

the centripetal force needs to be greater for a tighter turn (one with a smaller radius) than a broader one (one with a larger radius). On a level surface, side friction serves as a countering force to the centrifugal force, but it generally provides very little resistance/force. Thus, a vehicle has to make a very wide circle in order to make a turn on the level.

Given that road designs usually are limited by very narrow design areas, wide turns are generally discouraged. To deal with this issue, designers of horizontal curves incorporate roads that are sloped at a slight angle. This slope is defined as super elevation, which is the amount of rise seen on an angled cross-section of a road given a certain run. The presence of super elevation on a curve allows some of the centripetal force to be countered by the ground, thus allowing the turn to be executed at a faster rate than would be allowed on a flat surface. Super elevation also plays another important role by aiding in drainage during precipitation events, as water runs off the road rather than collecting on it. Generally, super elevation is limited to being less than 14 percent, as engineers need to account for stopped vehicles on the curve, where centripetal force is not present.

The radius of horizontal curve is also an important design aspect of the geometric design and road safety. The allowable radius (R) for a horizontal curve can be determined by knowing the design velocity (V), the coefficient of side-friction factor (f_s), gravitational force (g) and the allowable super elevation on the curve (e).

Radius of horizontal curve, $R = \frac{V^2}{g(e+f_s)}$ (3.2)

The radius of the curve determines flatness and sharpness of it. The smaller radius of the curve, the sharper the curve will be. For high speed highway, curves must be flat and have large radius to prevent unnecessary traffic accidents.

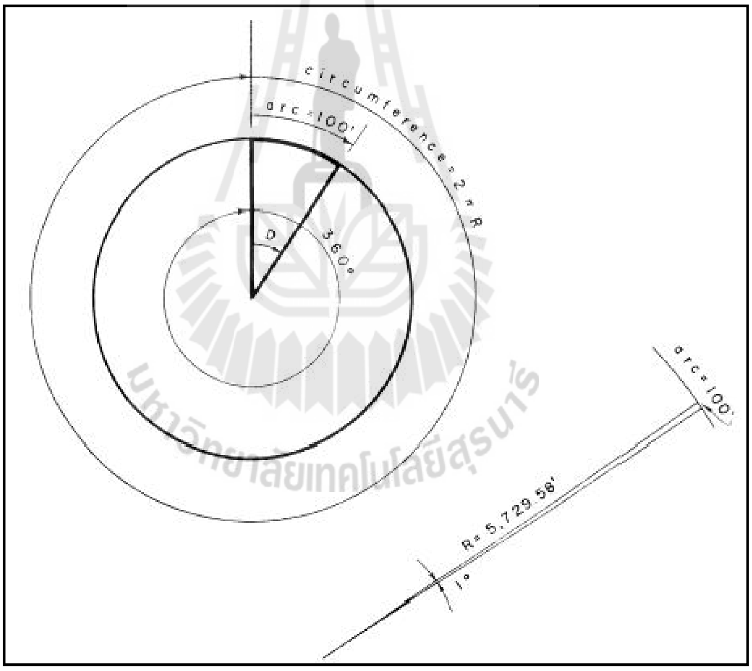
Accidents on horizontal curve are a case of concern in all countries, whatever the level of development of their road system (Elvik and Muskaus, 1994). While driving on country roads; driver's perceptual expectation of the trajectory of the road mainly depends on the road geometric alignment. When the road is mainly straight, drivers do not always expect sudden sharp curves to occur. When the road has numerous curves, on the other hand, drivers are more likely to expect further curves on the road ahead. Accordingly, higher accident rates have been found in sharp curves on roads with only few curves, compared to sharp curves on road with many sharp curves.

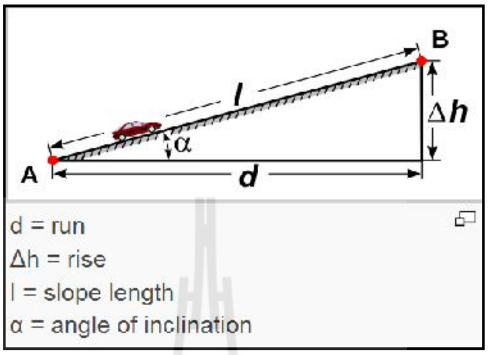
3.4.3 Degree of Curvature

The degree of curvature is defined as the central angle to the ends of an arc or chord of agreed length. The chord definition, displaying in fig 3.4, is used in railway practice and in some highway work. This definition states that the degree of curve is the central angle formed by two radii drawn from the center of the circle to the ends of a chord 100 feet (or 100 meters) long.

The ratio between the degree of curvature (D) and 360° is the same as the ratio between 100 feet of arc and the circumference (C) of a circle having the same radius. The relationship between horizontal curve radius degrees of curvature can be expressed as follows:

$$\frac{D}{360^{\circ}} = \frac{100}{C} \qquad (3.3)$$


circumference = 2 π R
arc = 100'
D
360°
R = 5,729.58'
1°
arc = 100'



$$\frac{Rise}{Run}$$

$$\frac{Rise}{Run}x100$$

driver’s ability to control the vehicle. The two important accepts are; the grade needs to be compatible with the braking capabilities of the vehicle and the grade will affect a vehicle’s stopping distance.

3.4.5 Combination of Horizontal and Vertical Alignment

Horizontal and vertical alignments are permanent design for study justification. It is extremely difficult and costly to correct alignment deficiencies after a highway is constructed. When manipulating a combination of horizontal and vertical alignment, it is important to consider the effects of the combination of both in the design stage. The combination of horizontal and vertical alignment produces the direct effect on the driver’s perception, operating speed, driver’s visual demand, vehicle stability, aesthetics of highway and safety. According to the study of Smith and Lam (1994), overlapping sag vertical curve would cause a horizontal curve to appear flatter while an overlapping crest curve would cause a horizontal curve to look sharper. This finding may reflect the driver to underestimate or overestimate on their perception. Regarding this finding, field measurements were carried out by Hassan et al., (2003). They tried to make sure the findings about the drivers' perception of the combinations of horizontal and vertical alignments. 1211 speed observations were collected on 6 sites of combinations of horizontal and sag curves and 1329 observations on 7 sites of combinations of horizontal and crest curves. And they confirmed that driver behavior on the approach to the horizontal curve varies with the type of overlapping vertical curve. Drivers commonly decelerated their operating speeds on the approach to crest combinations while drivers accelerated just before the beginning of horizontal curves in sag combinations. These finding may reflect the driver to underestimate or overestimate on their perception at the location where horizontal and vertical alignment combination.

3.4.6 Presence of Bridge along the Expressway

Bridges are vital components of the transportation infrastructure in all countries. A highway bridge carries a highway over an obstruction. Highway bridge traffic safety can be affected by many factors. These factors are road way width, shoulder width, approach guard rail and bridge rail, approach sight distance, approach roadway width, approach road way curvature, approach gradient, lighting, signing and delineation, presence of nearby ramps or intersections and presence of nearby lane drops or pavement transitions etc.

3.4.7 Presence of village settlement

Increasing human settlement adjacent to an expressway has conflicts between traffic movement and access of local people to that expressway and this is challenging road safety issue. To improve safety, highways should be designed to by-pass village settlements if it is possible.

3.5 General review for basic Crash Prediction Model

An accident prediction model can be used by jurisdictions to make better safety decisions and as part of network screening to identify sections that may have the best potential for improvements**.** The basic form of accident prediction models is as follow:

$$E(y) = \alpha \times \text{Traffic Volume}^{\beta_1} \times \text{Section Length}^{\beta_2} \quad (3.9)$$

The estimated expected number of accidents, E(y), is a function of traffic volume and section length. The effect of traffic volume and section length on accident frequencies is modeled in terms of an elasticity that is a power, β, to which traffic volume is raised.

The effects of other various risk factors can also influence the probability of accidents; the risk factors may concern the road geometric, road environment and road user behavior. Therefore, a complex accident prediction model can be expressed as the following equation.

$$E(y) = \alpha \times \text{Traffic Volume}^{\beta_1} \times \text{Section Length}^{\beta_2} \times \sum_{i=0}^{n} \beta_i X_i \tag{3.10}$$

Where; $E(y)$ = predicted number of accident

X_i = set of variable denoting as risk factors

α, β = model intercept and coefficients

The traffic volume and risk factors are the independent variables of the model. The choice of these variables to be included in an accident prediction model ought to be based on theory. However, the usual basis for choosing explanatory variables appears to be simply data availability. They should include variables that:

- have been found in previous studies to exert a major influence on the number of accidents;
- can be measured in a valid and reliable way;
- are not very highly correlated with other explanatory variables included.

It is recommended that basic APMs are developed for several road types depending on the national situation. Basic means that no risk factors are included, and only the traffic volume is used. In general, the accident numbers will be higher at increasing traffic volumes, but the accident rate will drop. These APMs could be used to benchmark one's roads. If the expected number of accidents is significantly lower than what is observed in reality, it is likely that there are some flaws in road design (Eenink et al., 2008). A good and detailed APM requires much data of good quality and detail that is usually not available. The APM outcomes can differ in different regions or countries.

3.6 Statistical Overview of Crash Prediction Model

The study mainly focused on the relationships between traffic crash and road characteristics to predict the number of crash. It is generally known that there is a complex relationship between traffic crash frequency and road characteristics. A factor or a combination of factors is an origin of crash occurrences. To control this complexity, crash prediction models can be used as a tool. Crash prediction models are used to predict the number of crashes on highways based on various crash modification factors such as traffic volume, lane width, shoulder width, degree of curvature etc.

In statistical modeling, regression analysis is a statistical process for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent (response) variable and one or more independent (predictors) variables. More specifically, regression analysis helps to understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed. The two basic type of regression model are linear and non linear model.

3.6.1 Traditional Linear Model

To review accident prediction model, traditional linear model is the first introduce among other alternative statistical regression models. A traditional linear model is of the form;

$$Y_i = X'i\,\beta + \varepsilon\ I \tag{3.11}$$

Yi = response variable of the i^{th} observation

Xi = explanatory variables of the i^{th} observation

The vector of unknown parameter β is estimated by a least squares fit to the data Y, which is normally distributed. The εi are assumed to be independent, normally distributed with zero mean and constant variance, εi~N (0, σ^2). X is fixed and constant variance σ^2·The expected value of Yi, denoted by μi = X ′i β.

Traditional linear models have types of problem while these are used extensively in statistical data analysis; the problems are as follow:

- It might not be reasonable to assume that data are normally distributed. For example, the normal distribution (which is continuous) might not be adequate for modeling counts or measured proportions that are considered to be discrete.
- If the mean of the data is naturally restricted to a range of values, the traditional linear model might not be appropriate, since the linear predictor $x_i'\beta$ can take on any value. For example, the mean of a measured proportion is between 0 and 1, but the linear predictor of the mean in a traditional linear model is not restricted to this range.

➢ It might not be realistic to assume that the variance of the data is constant for all observations. For example, it is not unusual to observe data where the variance increases with the mean of the data.

3.6.2 Generalized Linear Model

A generalized linear model is the extension of the traditional linear model and is therefore applicable to a wider range of data analysis problems. In statistics, the generalized linear model (GLM) is a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution. There are three components to any GLM:

➢ *Random Component* – refers to the probability distribution of the response variable (Y); e.g. normal distribution for Y in the linear regression, or binomial distribution for Y in the binary logistic regression, it is also called a noise model or error model.

➢ *Systematic Component* - specifies the explanatory variables (X_1, X_2, ... X_k) in the model, more specifically their linear combination in creating the so called linear predictor; e.g., $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots + \beta_n x_{2n}$.

➢ *Link Function*, η *or* g (μ) - specifies the link between random and systematic components. It says how the expected value of the response relates to the linear predictor of explanatory variables; e.g., $\eta = g\ (E\ (Y_i)) = E\ (Y_i)$ for linear regression, or η = *logit* (π) for logistic regression.

Assumptions:

❖ The data Y_1, Y_2... Y_n are independently distributed, i.e., cases are independent.

❖ The dependent variable Y_i does NOT need to be normally distributed, but it typically assumes a distribution from an exponential family (e.g. binomial, Poisson, multinomial, normal...)

❖ GLM does NOT assume a linear relationship between the dependent variable and the independent variables, but it does assume linear relationship between the transformed response in terms of the link function and the explanatory variables; e.g., for binary logistic regression $logit(\pi) = \beta_0 + \beta X$.

❖ Independent (explanatory) variables can be even the power terms or some other nonlinear transformations of the original independent variables.

❖ The homogeneity of variance does NOT need to be satisfied. In fact, it is not even possible in many cases given the model structure, and *over dispersion* (when the observed variance is larger than what the model assumes) maybe present.

❖ Errors need to be independent but NOT normally distributed.

❖ It uses maximum likelihood estimation (MLE) rather than ordinary least squares (OLS) to estimate the parameters, and thus relies on large-sample approximations.

❖ Goodness-of-fit measures rely on sufficiently large samples, where a heuristic rule is that not more than 20% of the expected cells counts are less than 5.

3.6.2.1 Poisson Regression Model

The Poisson regression model, a family of generalized linear model, is a technique used to describe count data as a function of a set of predictor variables. Poisson distribution focuses on the number of discrete events or occurrences over a specified interval, (time, length, distance etc) and can be expressed as follows;

$$\lambda = \frac{\text{number of occurrences}}{\text{specified interval}} \tag{3.12}$$

E(x) = expected value = $\mu = \lambda$,

Poisson regression model assumes that the dependent variable yi, the number of accidents in i^{th} highway segments during a period of time, follows the Poisson distribution with a parameter λi which is the expected accident frequency (or number of accidents) for i^{th} highway section during a period of time. The probability distribution of the model is as follows;

$$\Pr\left(Y=\frac{yi}{\lambda i}\right) = \frac{e^{-\lambda}\ \lambda^{yi}}{yi!}\ ,\ y_i = 0, 1, 2\ldots \tag{3.13}$$

λ is the mean or expected value of Poisson distribution.

- λ is also the variance of a Poisson distribution.
- Poisson has one parameter λ (lambda).

Pr (Y) in equation 14 is the probability of y accidents occurring at i^{th} highway section during a period of time. In Poisson regression model, the expected accident frequency is assumed to be a function of explanatory variables such that:

$$\text{Log}\ (\lambda i) = (\beta_0+\beta_1 * X_{i1}+\beta_2 * X_{i2}+\ldots+\beta_n * X_{in}) \tag{3.14}$$

Where, X_{i1}, X_{i2},…, X_{in} are the explanatory variables which include the traffic and highway characteristics of i^{th} highway section. The model coefficients, β_0, β1, β2,…,β_n , are estimated by maximum likelihood method. The likelihood function is as follows:

$$L\left(\frac{\beta}{y,X}\right) = \prod_{i=1}^{n} \Pr\left(\frac{yi}{\mu i}\right) = \prod_{i=1}^{n} * \frac{e^{-\mu i} \ \mu i^{yi}}{yi^{!}} \tag{3.15}$$

Where $\mu i = E\left(\frac{yi}{xi}\right) = \exp(x_{i*}\beta)$

Poisson distribution has a strong assumption of mean equal to variance, which is not generally exhibited by the crash data.

3.6.2.2 Negative Binomial Regression Model

Negative Binomial regression is a type of generalized linear model and it is the simple extension of the regular Poisson Regression to allow the variance to differ from its means. Therefore, an error structure is added to the Poisson mean parameter assumed to follow gamma distribution with mean 1 and dispersion parameter α. NB regression can overcome possible over dispersion parameter. The probability function of the model is as follows;

$$\Pr\left(Y = \frac{yi}{\lambda,\alpha}\right) = \frac{\Gamma\left(yi+\alpha^{-1}\right)}{y!\,\Gamma\left(\alpha^{-1}\right)} \left(\frac{\alpha^{-1}}{\alpha^{-1}+\lambda}\right)^{\alpha^{-1}} \left(\frac{\lambda}{\alpha^{-1}+\lambda}\right)^{yi}, = 1, 2, 3\ldots \tag{3.16}$$

- The negative binomial distribution has two parameters: λ and α.
- λ is the mean or expected value of the distribution.
- Γ is a value of gamma distribution.

The symbol (α) in equation 17 can be denoted as a measure of dispersion or over dispersion parameter. When α becomes zero, the negative binomial distribution is the same as a Poisson distribution. In Negative Binomial regression model, the expected accident frequency is assumed to be a function of explanatory variables such that:

$$\text{Log}(\lambda i) = (\beta_0+\beta_1 * X_{i1}+\beta_2 * X_{i2}+\ldots+\beta_n * X_{in}) \quad (3.17)$$

Where, X_{i1}, X_{i2},…, X_{in} are the explanatory variables which include the traffic and highway characteristics of i^{th} highway section. $\beta_0,\beta_1,\beta_2,\ldots,\beta_n$ are the model coefficients estimated by maximum likelihood method. The likelihood function is as follows:

$$L\left(\frac{\beta}{y,X}\right) = \prod_{i=0}^{n} \Pr\left(\frac{yi}{xi}\right) = \prod_{i=0}^{n} \frac{\Gamma(y+\alpha^{-1})}{y!\,\Gamma(\alpha^{-1})} \left[\frac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right]^{\alpha^{-1}} \left[\frac{\mu_i}{\alpha^{-1}+\mu_i}\right]^{y_i} \quad (3.18)$$

Where, $\mu_i = E\left(\frac{y_i}{x_i}\right) = \exp(x_i * \beta)$

3.6.3 Correlation Definition and Assumption

Correlation is a bivariate analysis that measures the strength of association between two variables. In statistics, the value of the correlation coefficient varies between +1 and -1. When the value of the correlation coefficient lies around ± 1, then it is said to be a perfect degree of association between the two variables. As the correlation coefficient value goes towards 0, the relationship between the two variables will be weaker. Generally, there are three types of correlations in statistics; Pearson correlation, Kendall rank correlation and Spearman correlation.

3.6.3.1 Pearson's Product-Moment Correlation

The Pearson product-moment correlation coefficient (Pearson’s correlation, for short) is a measure of the strength and direction of association that exists between two variables measured on at least an interval scale. A Pearson’s correlation attempts to draw a line of best fit through the data of two variables, and the Pearson correlation coefficient, *r*, indicates how far away all these data points are to this line of best fit (i.e., how well the data points fit this new model/line of best fit). The following formula is used to calculate the Pearson correlation:

$$r = \frac{N\sum XY - \sum(X)(Y)}{\sqrt{[N\sum X^2 - \sum(X)^2][N\sum Y^2 - \sum(Y)^2]}} \tag{3.19}$$

Where: r = Pearson r correlation coefficient

N = number of value in each data set

$\sum XY =$ Sum of the products of paired scores

$\sum X =$ Sum of x scores

$\sum Y =$ Sum of y scores

$\sum X^2 =$ Sum of squared x scores

$\sum Y^2 =$ Sum of squared y scores

Assumptions

- Variables should be measured at the interval or ratio level (i.e., they are continuous).
- There needs to be a linear relationship between the two variables.
- There should be no significant outliers.
- Variables should be approximately normally distributed

3.6.3.2 Spearman's Rank-Order Correlation

The Spearman rank-order correlation coefficient (Spearman's correlation, for short) is a non-parametric measure of the strength and direction of association that exists between two variables measured on at least an ordinal scale. It is denoted by the symbol r_s (or the Greek letter ρ, pronounced rho). The test is used for either ordinal variables or for continuous data that has failed the assumptions necessary for conducting the Pearson's product-moment correlation. The following formula is used to calculate the Spearman rank correlation:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3.20}$$

Where: ρ = Spearman rank correlation

d_i = the difference between the ranks of corresponding values Xi and Yi

n = number of value in each data set

Assumptions

- Variables should be measured on an ordinal, interval or ratio scale.
- There needs to be a monotonic relationship between the two variables. A monotonic relationship exists when either the variables increase in value together, or as one variable value increases, the other variable value decreases.

3.6.3.3 Kendall rank correlation

Kendall rank correlation is a non-parametric test that measures the strength of dependence between two variables. The following formula is used to calculate the value of Kendall rank correlation:

$$\tau = \frac{n_c - n_d}{\frac{n(n-1)}{2}} \tag{3.21}$$

Where: τ = Kendall rank correlation coefficient

$n_{c=}$ Number of concordant

$n_{d=}$ Number of discordant

3.7 Empirical Bayesian (EB) Estimation Method

Empirical Bayesian (EB) Method was developed by Reverend Thomas Bayes and it can increases the precision of estimation and corrects for the regression-to-mean bias. The EB Method adjusts the determination of expected number of accident which is output from accident prediction model. The EB approach is recommended as a state-of-the-art approach to estimate expected number of crashes and identifying hazardous highway locations. The EB Method is applicable at either the site-specific level (where crashes can be assigned to a particular location) or the project specific level (where observed data may be known for a particular facility, but cannot be assigned to the site specific level). Where only a predicted or only observed crash data are available, the EB Method is not applicable (Highway Safety Manual, 2010).

3.7.1 The EB Procedure

The procedure combines the recorded crash number of every segment and the expected crash frequency by the prediction model for similar segments. The basic functional form is as follows;

$$N_{(expected)} = w \times N_{(predicted)} + (1-w) \times N_{(observed} \tag{3.22}$$

Where: $N_{(expected)}$ = expected average crashes by EB estimate

$N_{(predicted)}$ = predicted average crashes estimated by prediction model

$N_{(observed)}$ = observed or recorded crash frequency

w = weighting adjustment factor

The weighted adjustment factor, w, is a function of the crash prediction model's over dispersion parameter, k, and is calculated using following equation.

$$w = \frac{1}{1+k\sum_{i=0}^{n} N\ (predicted)} \tag{3.23}$$

Where: k= dispersion parameter from crash prediction model.

When the value of the over dispersion parameter increases, the value of the weighted adjustment factor will decrease. Hence, more emphasis is placed on the observed rather than the predicted crash frequency.

3.8 Methods for Identifying Hazardous Road Locations

Identifying hazardous road locations is the first step of road safety management program and can be carried out in many different ways. Hazardous location can be defined as the one that has a higher expected number of accidents than similar locations due to local risk factors. Identifying hazardous location is great important to land transport authorities and other decision makers to logically allocate the budgets as well as providing other resources in a cost-effective reduction solution.

Zeeger (1982), Gharaybeh (1991),Utainarumol (1999) and Ratanavaraha et.,al (2003) applied five methods for identification of hazardous locations. These were (1) Accident Frequency Method, (2) Accident Rate Method, (3) Quality Control Method, (4) Accident Severity Method and (5) Combined Method. In this study, only four methods (Accident Frequency Method, Accident Rate Method, Quality Control Method and Combined Method) were performed. Accident Severity Method was not considered in the study because of unreliable accident severity data.

3.8.1 Accident Frequency Method

This method is used for exploration of concentration of accidents within a fixed or variable segment length. The accident frequencies of each segment are ranked in terms of descending order and the segment having highest number of accidents was determined as first hazardous location.

3.8.2 Accident Rate Method

In this method, accident rate was obtained by dividing accident frequency for given analysis period and the million vehicle-miles of travel at highway segment. And then, segments were ranked in descending order. The segment with highest accident rate was defined as first hazardous location.

$$R = \frac{A*1,000,000}{365*T*V*L} \tag{3.24}$$

where: R =accident rate for highway segment (in accidents per million vehicle miles),

A =number of accidents for given analysis period,

T = time of analysis period (in years or fraction of years),

V = average annual daily traffic (AADT) during study period, and

L = length of highway segment (in miles).

3.8.3 Rate Quality Control Method

In Rate Quality Control Method, the accident rate, output of Accident Rate Method, was used to calculate the average accident rate (Ra) for all segments. This average accident rate (Ra) was then input in the equation to find critical accident rate (Rc) for highway segment.

$$Rc = Ra + K(Ra/E)^{0.5} + 1/(2E) \tag{3.25}$$

Where: Rc =critical accident rate for highway segment (accidents per million vehicle-miles)

Ra = average accident rate for all highway segments of similar characteristics

E = (365*T*V*L)/1,000,000, million vehicle-miles of travel on the highway segment during the study period and

K = a probability factor determined by the desired level of significance for the equation. K is representing 95% confidence level and its value is 1.645.

The locations are ranked based on the ratio of actual accident rate (R) and critical accident rate (Rc). The greater (R/Rc) value was determined to be more dangerous and ranked first.

3.8.4 Combined Method

This method is combined by the previous three methods (Accident Frequency Method, Accident Rate Method and Quality Control Method). The result was defined as danger index (DI) and it was obtained by taking sum of the rank numbers of previous three methods. Therefore the location that has a lower value of DI can be defined as more hazardous and it was ranked first.

$$DI = (F\_Rank + R\_Rank + Q\_Rank) \qquad (3.26)$$

Where; DI = Danger Index

F_Rank = rank of location by accident frequency method

R_Rank = rank of location by accident rate method

Q_Rank = rank of location by quality control method

Individual methods such as accident frequency, accident rate and rate quality control are not suitable to be used alone to identify hazardous highway locations because the results from these individual methods are so highly dependent upon input data. Rankings can drastically change when some situations are changed, such as changing amount of data and time periods, and sometimes the method presents unclear results (Utainarumol and Stammer, 1999).

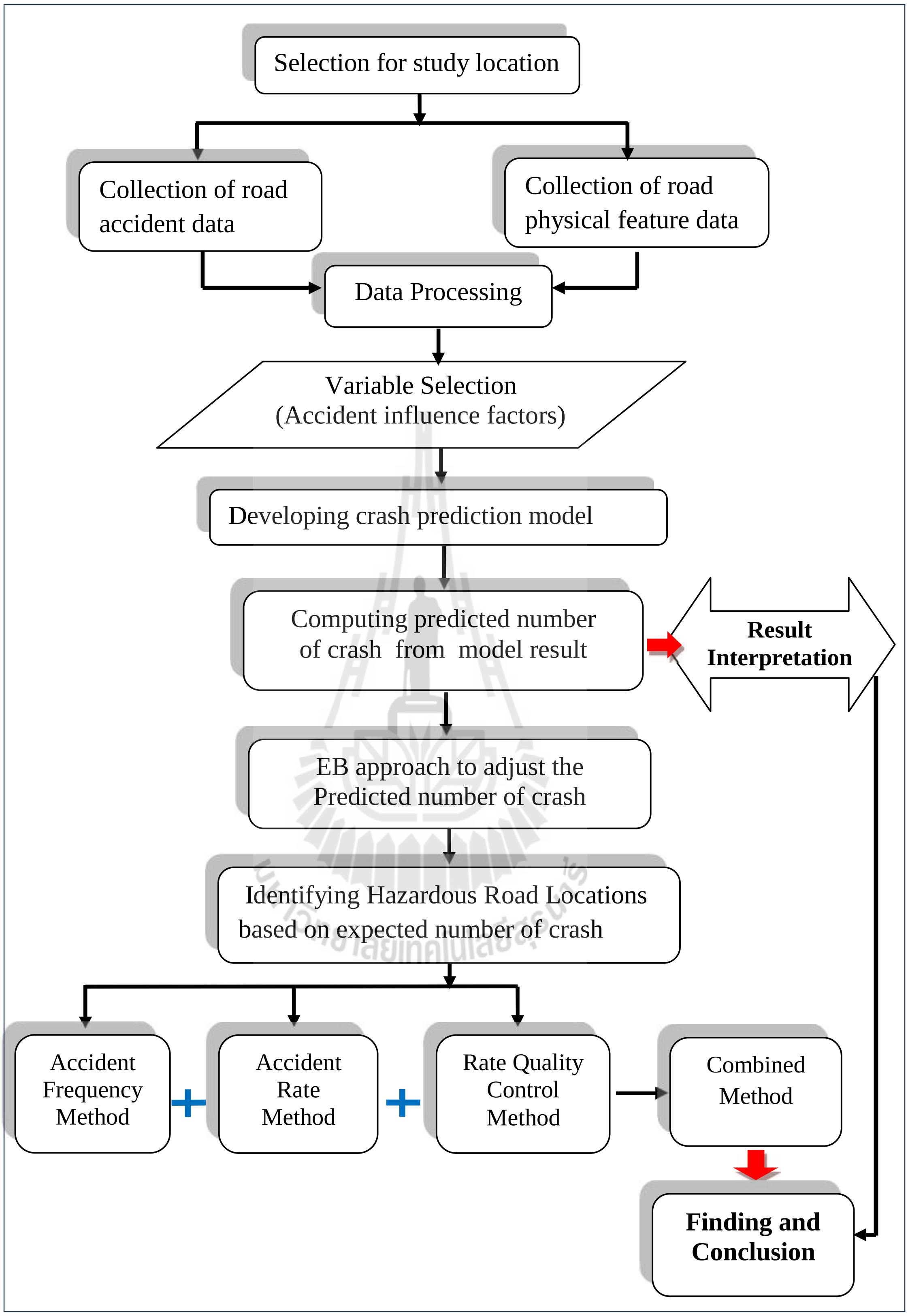


Figure 3.6 Flow chart of methodology

3.9 Summary of the Methodology

The aim of the study was to identify hazardous highway locations and to investigate the relationship between accident frequency and road characteristics on Yangon-Mandalay Expressway. Figure 3.6 indicates the method which was applied in the study and consists of nine steps. First, a total of 366mile (589km) long Yangon-Mandalay Expressway was considered as study location. Second, required data were collected from Ministry of Construction. The data comprised two types, road accident data and road physical feature data. These data were then extracted and processed by using Microsoft Excel application. In third step, variables that approximately influence on accident occurrences were selected based on current road condition. After considering variable selection, an accident prediction model (Negative Binomial Regression) was developed in the fourth step. The fifth step was modal result interpretation and it was the finding of relationship between accident frequency and road characteristics on Yangon-Mandalay Expressway. In the sixth step, expected numbers of accident were computed based on the result of crash prediction model that had taken in the fifth step. The main idea of this step was for identifying hazardous road locations based on expected number of crash rather than observed number. Regarding the previous study of identifying hazardous road locations, it was noted that the empirical Byes (EB) estimation method was the best suited for obtaining the expected number of accidents and the use of it should rely on an accident prediction model (Elvik 2008). Therefore EB estimation method was performed in the seventh step. Applying empirical Byes (EB) method enhanced precision of the estimation and eliminates regression to the mean bias in the crash data. After obtaining expected number of accident adjusted by EB estimation method, in the eighth step, the four methods (Accident Frequency Method, Accident Rate Method, Quality Control Method and Combined Method) were carried out in order to identify hazardous locations on Yangon-Mandalay expressway. Summarizing finding and making conclusion for the study were accomplished at the final step.

CHAPTER IV

ANALYSIS AND RESULT

4.1 Introduction

This chapter focuses on the analysis of the study. The purposes of the study were (1) identifying hazardous highway locations based on the expected number of accident and (2) examining relationships between crash frequency and road characteristics. Expected numbers of accident on each of road segments were predicted based on the road characteristics variable by developing an Accident Prediction Model (APM). The success of traffic accident prediction model depends on the quantity and quality of the data used to calibrate the model. And there is no precise answer regarding the amount of data required for a good model to be calibrated.

The statistical software package, SPSS (Statistical Package for the Social Sciences), was used for calibrating accident prediction model for the study. This is a widely used program for statistical analysis in social science and can perform highly complex data manipulation and analysis with simple instructions. In this section, Negative Binomial Regression model was performed followed by goodness of fit measures. The goodness of fit of a statistical model describes how well a model fits the observed data. In other words, measures of goodness of fit typically summarize the variation between observed values and the values expected under the model in question.

4.2 Defining Study Area

Yangon-Mandalay expressway, whose total length was 366 miles, was defined as study location. Road sections were divided into one mile fixed length segment to calibrate the accident prediction model. Hence a total of 732 observations were obtained for both directions of travel.

4.3 Variable Description and Ranking

The latest three years traffic crash data (from 2013 to 2015) were set up as response variable. The other road characteristics variables such as Average Daily Traffic, Presence of bridge, Presence of sharp horizontal curve, Average horizontal curvature, Percent downgrade Percent upgrade, Combination of horizontal curve and slope and Presence of village settlement were handled as predictor variables in the crash prediction model. Table 4-1 displays the selected variables for calibrating crash prediction model and these variables were specified according to their type and nature.

4.4 Development of Crash Prediction Model

4.4.1 Descriptive statistic

Descriptive statistics, in short, help describe and understand the features of a specific data set, by giving short summaries about the sample and measures of the data. To summarize the data set of the study, Descriptive statistics were carried out and these were broken down into measures of central tendency and measures of variability, or spread. Measures of central tendency included the mean, while measures of variability included the standard deviation or variance and the minimum and maximum variables. The descriptive statistics for the study is shown in table 4.2. The value of variance and mean of the dependent variable (number of accident) were (2.59) and (1.33) respectively. This means that conditional variance of dependant variable is greater than its conditional mean, which we call over dispersion phenomena in crash data. To overcome this complexity, Negative Binomial regression model was performed with the help of SPSS software.

Table 4.1 Variable Selection and description

No	Classifying variable for analysis	Variable description	Type of Variable
1	Response variable	Number of accident in three years per road segment	Continuous variable
2	Predictor variables	Annual Average Daily Traffic per road segment	Continuous variable
3		Presence of sharp horizontal curve per road segment	Categorical variable 1= present ,0 = absent
4		Average horizontal curvature per road segment	Continuous variable
5		Percent upgrade per road segment	Continuous variable
6		Percent downgrade per road segment	Continuous variable
7		Combination of horizontal curve and slope per road segment	Categorical variable 1= yes, 0 = otherwise
8		Presence of bridge per road segment	Categorical variable 1= None 2= one bridge presence 3= two and more bridges presence
9		Presence of village settlement per road segment	Categorical variable 1= present ,0 = absent

Table 4.2 Descriptive statistic of response and predictor variables

	N	Min;	Max;	Mean	Std. Deviation	Variance
Number of accident	732	0	21	1.33	1.61	2.59
Ln -AADT	732	7.15	8.06	7.63	.24	.060
Presence of sharp horizontal curve	732	0	1	.47	.50	.250
Average horizontal curvature	732	0	146	6.61	11.92	142.14
Percent upgrade	732	0	5	.22	.448	.20
Percent downgrade	732	0	5	.33	.548	.30
Combination of horizontal curve and slope	732	0	1	.39	.488	.23
Presence of bridge	732	1	3	1.67	.72	.52
Presence of village settlement	732	0	1	.20	.40	.16
Valid N (list wise)	732					

4.4.2 Correlation Test

A number of tests were carried out between variables used in the prediction model to check whether they met the requirements of correlation assumptions or not. And it was found that there were some consequences associated with a violation of the Pearson's correlation assumption in the data. Therefore Spearman rank correlation analysis was accomplished to figure out correlations between variables; not only between response variable and predictor variables but also between predictor variables itself. Low correlation coefficient showed that there was not a strong association between these variables and high correlation coefficient showed that high association exists between them. According to the statistical theory, the value of correlation coefficient between predictor variables shall be low and it shall be high between response variable and predictor variables. Hence variables with high degree of correlation between predictor variables were removed from the model in order to obtain the better result.

In this regard, it was determined as a strong correlation if the value of coefficient was greater than 0.5. And it was also determined as weak correlation if the value of coefficient was 0.5 and less than 0.5. Result of correlation analysis of the study displays in table 4.3.

Table 4.3 Correlation Test Result

Correlations[c]

			No .of. Accident	Ln AADT	Presence .of . Sharp. Horizontal curve	Average . horizontal. Curvature	Percent . upgrade	Percent . downgrade	Combination .of. Horizontal curve. and. Slope	Presence .of. bridge	Presence .of. village . settlement
Spearman's rho	No .of. Accident	Correlation Coefficient	1.000	.200**	.237**	.091*	-.105**	.140**	.294**	.119**	.138**
		Sig. (2-tailed)	.	.000	.000	.014	.005	.000	.000	.001	.000
	Ln AADT	Correlation Coefficient	.200**	1.000	.389**	.395**	-.005	.039	.312**	.096**	.059
		Sig. (2-tailed)	.000	.	.000	.000	.899	.296	.000	.010	.113
	Presence .of .Sharp. Horizontal curve	Correlation Coefficient	.237**	.389**	1.000	.769**	-.001	.057	.621**	.155**	.036
		Sig. (2-tailed)	.000	.000	.	.000	.969	.122	.000	.000	.334
	Average .horizontal. Curvature	Correlation Coefficient	.091*	.395**	.769**	1.000	.021	.094*	.590**	.164**	.046
		Sig. (2-tailed)	.014	.000	.000	.	.568	.011	.000	.000	.210
	Percent .upgrade	Correlation Coefficient	-.105**	-.005	-.001	.021	1.000	-.630**	.093*	-.037	-.071
		Sig. (2-tailed)	.005	.899	.969	.568	.	.000	.011	.316	.056
	Percent .downgrade	Correlation Coefficient	.140**	.039	.057	.094*	-.630**	1.000	.187**	-.018	.016
		Sig. (2-tailed)	.000	.296	.122	.011	.000	.	.000	.626	.656
	Combination .of. Horizontal curve. and. Slope	Correlation Coefficient	.294**	.312**	.621**	.590**	.093*	.187**	1.000	.123**	.065
		Sig. (2-tailed)	.000	.000	.000	.000	.011	.000	.	.001	.077
	Presence .of. bridge	Correlation Coefficient	.119**	.096**	.155**	.164**	-.037	-.018	.123**	1.000	.068
		Sig. (2-tailed)	.001	.010	.000	.000	.316	.626	.001	.	.065
	Presence .of. village . settlement	Correlation Coefficient	.138**	.059	.036	.046	-.071	.016	.065	.068	1.000
		Sig. (2-tailed)	.000	.113	.334	.210	.056	.656	.077	.065	.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

c. Listwise N = 732

4.4.3 Modeling

Base on the previous literature surveys, a basic Safety Performance Function (SPF), referred to Crash Prediction Model (CPM), always includes traffic volume (AADT) and road segment length (L). These variables can be denoted as exposure variables since the predicted number of crash per mile/km is a function of AADT and road segment length. But other site characteristics variables such as lane width, shoulder width, radius/degree of horizontal curves, and presence of turn lanes may also include in a complex SPF.

According to the descriptive statistic of the study, the mean and variance of crash data ware (1.33) and (2.59) respectively. Hence, the frequency of crashes was assumed to follow a Negative Binomial distribution as variance of response variable was greater than its mean. And this was consistent with previous research papers (Dissanayake and Ratnayake, 2006), (Kibar et al., 2013) and (Naznin et al., 2016). Negative Binomial regression is a type of generalized linear model and it is the simple extension of the regular Poisson Regression to allow the variance to differ from its means. The link function of the model is log and model coefficients are estimated by the maximum likelihood approach. The relationship between expected numbers of crash (Y_i) and predictor variables (X_{i1}, X_{i2}… X_{in}) related to road characteristics on Yangon-Mandalay expressway can be expressed as:

$$\ln E(Y) = \beta_0 \times \beta_1 \ln \text{AADT} \times (\textstyle\sum_{i=0}^{n} \beta_2 X_2 + \beta_3 X_3 + \cdots \beta_i X_i\ ) \tag{4.1}$$

Where; E(Y) = predicted number of crash on a segment length

AADT = Annual Average Daily Traffic

X_2 = Presence of bridge per road segment

X_3 = Presence of sharp horizontal curve per road segment

X_4 = Average horizontal curvature per road segment

X_5 = Percent downgrade per road segment

X_6 = Percent upgrade per road segment

X_7 = Combination of horizontal curve and slope per road segment

X_8 = Presence of village settlement per road segment

The predictor variable, Annual Average Daily Traffic (AADT) referred to exposure variable in the model and it was transformed into natural log. Transforming exposure variable provided a better fit and was suitable for functional representation of larger AADT values, (Valentova et al., 2014). Furthermore, transforming exposure measures to the natural logarithm was common practice of crash prediction modeling in previous researches, (Ackaah and Salifu, 2011), (Ceunynck et al., 2011), and (Valentova et al., 2014).

4.4.4 Model Evaluation

Two statistical measures were used to evaluate the statistical performance of the model. These were the Pearson Chi-square statistics and the Deviance statistics. Table 4.4 displays the goodness of fit measures for the study. The value of Pearson Chi-square and Deviance statistics divided by its degree of freedom were estimated to be 1.081 and 1.088 respectively. This showed that the assumption of Negative Binomial (NB) distribution and the use of NB modal were appropriate for the data since these values were within the acceptable range, between 0.8 and 1.2 (Ackaah and Salifu, 2011; Dissanayake and Ratnayake, 2006). The value of dispersion parameter, estimated from the NB model which has shown in table 3, was found to be significantly different from zero ($\emptyset = 0.341$); this also indicated that the use of NB model was more suitable than using Poisson model (Naznin et al., 2016).

Table 4.4 Goodness of fit test of the model estimated by SPSS statistical software

Goodness of fit statistics	Value	df	Value/df
Deviance	787.360	724	1.088
Scaled Deviance	787.360	724	
Pearson Chi-Square	782.464	724	1.081
Scaled Pearson Chi-Square	782.464	724	
Log Likelihood[b]	-1102.3		
Akaike's Information Criterion (AIC)	2220.14		
Finite Sample Corrected AIC (AICC)	2220.34		
Bayesian Information Criterion (BIC)	2256.91		
Consistent AIC(CAIC)	2264.91		

4.4.5 Model Parameter Estimation

Table 4.5 represents the parameter estimation of the model result. Any insignificant variables ($p > 0.05$) throughout the process were omitted from the model and only significant variables ($p < 0.05$) were maintained in the model. According to the model result, five predictor variables including presence of bridge, Annual Average Daily Traffic (AADT), percent downgrade , combination of horizontal curve and slope and presence of village settlement were significant with statistical criterions, ($P < 0.05$ and 95% confident intervals for the coefficients). This means that crash frequency increases with increasing traffic volume, presence of bridge, increasing percent downgrade, combination of horizontal curve and slope, and presence of village settlement.

Table 4.5 Parameter Estimation of the Model estimated by SPSS statistical software

Parameter Estimates							
Parameter	B	Std. Error	95% Wald		Hypothesis Test		
			Lower	Upper	Wald	df	Sig.
(Intercept)	-4.859	1.353	-7.512	-2.206	12.88	1	.000
[Presence of bridge=3]	.087	.1165	-.141	.316	.560	1	.454
[Presence of bridge=2]	.275	.0862	.106	.444	10.19	1	.001
[Presence of bridge=1]	0[a]	.	.	.	.	.	.
Ln AADT	.609	.1780	.260	.958	11.69	1	.001
Percent downgrade	.242	.0635	.118	.367	14.57	1	.000
Combination of Horizontal curve and Slope	.422	.0840	.257	.587	25.22	1	.000
Presence of village settlement	.286	.0928	.104	.467	9.474	1	.002
(Scale)	1[b]						
(Negative binomial)	.341	.0602	.242	.482			

4.5 Result Interpretation for Crash Prediction Model

One of the predictor variables, presences of bridge was split into three levels and coded as (1 = none, 2 = one bridge presence and 3 = two or more bridges presence) in the modeling process. The model indicates that the presence of one bridge on the road segment correlates with a higher crash count. The other variables, no bridge presence and two or more bridges presence, on the road segment are insignificant. It implies that having more bridges on the same segment is less dangerous than having only one bridge. Drivers might be more aware of the presence of more bridges and therefore being more alert and drive more carefully. It has been noticed that some of the bridge on Yangon-Mandalay expressway have complex geometric approach, lack of bridge-approach guardrail with proper transition and end treatment. This situation might happen to bridge related crashes on the expressway.

The Annual Average Daily Traffic has been used as the exposure variable in all crash prediction models of previous researches and it has positive relationship with traffic accidents, (Saffarzadeh and Pooryari, 2005; Caliendo et al., 2006; Kumar, 2013). The result is compatible with these papers indicating that AADT is statistically significant with positive estimation coefficient; meaning that crash frequency increases with increasing traffic volume.

The variable describes percent downgrade in the model also correlates with higher crash frequency. This situation might concern with the skidding and demand of side friction on downgrade. Downgrade changes the distribution of weight on tires and they can consequently alter the dynamic performance of vehicles in terms of forces and accelerations. As a result, the side friction factor increases as the downgrade increases, (Kordani et al., 2014).

The variable naming combination of horizontal curve and slope in the model has a considerable influence on the crash frequency as indicated by the positive model parameter. The previous study has shown a similar result reporting that horizontal curves were more dangerous when combined with gradients and surface with low coefficients of friction (Aram, 2010). Horizontal and vertical curves should not be designed independently when a section of a highway needs to be designed with combined alignments. Imperfect combination of horizontal and vertical alignment may pose negative impact on driving comfort and worsen safety effect.

Ackaah and Salifu (2011) have found an increase of injury crashes in road sections which has village settlements compared with segments with no settlements.

Likewise, the predictor variable for presence of village settlement in the model displays that it positively correlates with crash count. Crash data used for this study include several types of crash such as collisions involving motorcycle and passenger vehicle, vehicle-animals collisions and vehicle-pedestrian collisions. The vast majority of these accidents can be related to the settlement of local people along the expressway.

4.6 Calculating Expected Number of Crash

4.6.1 Application of Model equation to find Predicted Number of Crash

According to the model result, crash frequency increases with increasing traffic volume, presence of bridge, increasing percent downgrade, combination of horizontal curve and slope, and presence of village settlement. Predicted number of crash for each road segment was calculated by following equation:

$$E(Y) = \mathrm{Exp}(-4.895 + 0.275PB + 0.609\ \ln AADT + 0.242PD + 0.422CHS + 0.286PV) \tag{4.2}$$

Where; E(Y) = Predicted numbers of crash per road segment

P B = Presence of bridge per road segment

AADT= Annual Average Daily Traffic per road segment

PD = Percent downgrade per road segment

CHS = Combination of horizontal curve and slope per road segment

PV = Presence of village settlement per road segment

Exp = Exponential function (2.718282)

Micro soft Excel application was utilized to calculate the predicted numbers of crash on each of road segments based on the model equation and table 4.6 represents the results for the first nineteen segments on Yangon-Mandalay expressway.

Table 4.6 Predicted numbers of crash estimated by model equation for the first nineteen segments on Yangon- Mandalay expressway

Begin; mile	End; mile	Observed number of Accident	Ln AADT	Percent downgrade	Combination of Horizontal curve and Slope	Presence of bridge	Presence of village settlement	Predicted no of accident
0	1	0	8.06	0.00	0	1	0	1
1	2	2	7.91	0.00	0	1	0	1
2	3	0	7.91	0.00	0	1	0	1
3	4	0	7.91	0.00	0	1	0	1
4	5	0	7.91	0.00	0	1	0	1
5	6	1	7.91	0.00	0	1	0	1
6	7	2	7.91	0.00	1	1	0	2
7	8	0	7.91	0.00	1	1	1	3
8	9	0	7.91	0.00	1	1	0	2
9	10	0	7.91	0.00	1	1	0	2
10	11	1	7.91	1.9	1	1	0	3
11	12	0	7.91	0.9	0	1	1	2
12	13	0	7.91	0.0	0	1	0	1
13	14	1	7.91	0.0	0	1	1	2
14	15	1	7.91	0.0	0	1	0	1
15	16	3	7.91	0.0	0	1	0	1
16	17	1	7.91	0.0	0	1	0	1
17	18	2	7.91	0.0	1	1	0	2
18	19	0	7.91	0.0	0	1	0	1

4.6.2 Weighting Predicted Numbers of Crash by Empirical Bayes (EB) Estimation Method

The EB estimation can increase the precision of estimation and corrects for the regression-to-mean bias. Hence the predicted number of crash on each of road segments that have been calculated above section was adjusted by EB procedure. This approach adjusts between observed numbers of crash on each road segments and predicted numbers of crash from the prediction model for similar segments. The following basic functional form was applied:

$$N_{expected} = w \times N_{predicted} + (1-w) \times N_{observed} \tag{4.3}$$

Where; $N_{(expected)}$ = expected numbers of crash by EB estimate

$N_{(predicted)}$ = predicted numbers of crash estimated by prediction model

$N_{(observed)}$ = observed or recorded numbers of crash

W = weighting adjustment factor, it was calculated as follow:

$$w = \frac{1}{1+k\sum_{i=0}^{n} N_{(predicted)}} \tag{4.4}$$

Where; k = over dispersion parameter from crash prediction model, (0.341fromNB model).

As the value of over dispersion parameter (k) increases, the value of weighted adjustment factor decreases. Thus, more emphasis is placed on the observed rather than predicted crash frequency. When the data used to develop a model are greatly dispersed, the reliability of the resulting predicted crash frequency is likely to be lower. In this case, it is reasonable to place less weight on the predicted crash frequency and more weight on the observed crash frequency. On the other hand, when the data used to develop a model have little over dispersion, the reliability of the resulting crash prediction model is likely to be higher. In this case, it is reasonable to place more weight on the predicted crash frequency and less weight on the observed crash frequency. Table 4.7 displays the standard deviation and coefficient of variation of observed crash frequency as well as expected crash frequency which is adjusted by EB method. It had been found out that expected crash frequency approached by EB adjustment has less variability than observed number of crash. Table 4.8 shows the calculation of expected number of crash adjusted by EB Method for the first nineteen segments on Yangon-Mandalay expressway.

Table 4.7 Descriptive Statistics for observed crash and expected crash by EB adjustment

	Mean	Std. Deviation	Coefficient of variation
Observed crash	1.33	1.611	1.21
Expected crash by EB estimation	0.86	0.405	0.47

Table 4.8 Expected number of accident estimated by EB Method for the first nineteen segments on Yangon-Mandalay expressway

Beginning mile	Ending mile	Observed no of crash	Predicted number of crash	w	N(expected)
0	1	0	1	0.54	1
1	2	2	1	0.44	1
2	3	0	1	0.44	1
3	4	0	1	0.44	1
4	5	0	1	0.44	1
5	6	1	1	0.44	1
6	7	2	2	0.34	1
7	8	0	3	0.28	1
8	9	0	2	0.34	1
9	10	0	2	0.34	1
10	11	1	3	0.24	1
11	12	0	2	0.32	1
12	13	0	1	0.44	1
13	14	1	2	0.37	1
14	15	1	1	0.44	1
15	16	3	1	0.44	1
16	17	1	1	0.44	1
17	18	2	2	0.34	1
18	19	0	1	0.44	1

4.7 Identification of Hazardous Highway Locations Analysis

Identifying hazardous road locations is the first step of road safety management program and can be carried out in different ways. For this study, the following methods were performed.

(1) Accident Frequency Method

(2) Accident Rate Method

(3) Rate Quality Control Method and

(4) Combined Method

The output from EB adjustment, expected numbers of crash, were utilized for identification of hazardous locations on Yangon-Mandalay expressway. In this regard, it was considered two directions of travel; north bound direction (Yangon-Mandalay) and south bound direction (Mandalay-Yangon). Microsoft Excel application was applied throughout the analysis process.

4.7.1 Accident Frequency Method

In this method, expected numbers of crash were allocated to the associated road segment and each of road segments has one mile interval. After allocating the expected numbers of crash, all of road segments were ranked in descending order. The segment with highest expected number of crash was determined as more hazardous and ranked first. Table 4.9 and 4.10 describe the first eighteen hazardous locations on Yangon-Mandalay expressway performed by Accident Frequency Method.

Table 4.9 Output of *Accident Frequency analysis* based on one mile interval for the most dangerous eighteen segments (*Yangon-Mandalay direction*)

Rank No	Beginning Mile	Ending Mile	N(expected)
1	125	126	6
2	113	114	3
3	84	85	3
4	275	276	2
5	39	40	2
6	101	102	2
7	63	64	2
8	75	76	2
9	105	106	2
10	358	359	2
11	242	243	2
12	92	93	2
13	276	277	2
14	68	69	2
15	185	186	2
16	114	115	2
17	104	105	2
18	83	84	2

Table 4.10 Output of *Accident Frequency analysis* based on one mile interval for the most dangerous eighteen segments (*Mandalay-Yangon direction*)

Rank No	Beginning mile	Ending Mile	N(expected)
1	115	116	2
2	83	84	2
3	117	118	2
4	89	90	2
5	64	65	2
6	67	68	2
7	27	28	2
8	116	117	2
9	60	61	2
10	129	130	2
11	184	185	2
12	8	9	2
13	352	353	2
14	125	126	2
15	32	33	2

Table 4.10 Output of *Accident Frequency analysis* based on one mile interval for the most dangerous eighteen segments (*Mandalay-Yangon direction*) (Cont.)

Rank No	Beginning mile	Ending Mile	N(expected)
16	58	59	2
17	103	104	2
18	3	4	1

4.7.2 Accident Rate Method

In Accident Rate Method, it was simply divided the expected numbers of crash on each of road segment by the vehicle exposure; determined as the number of accidents per million vehicle-miles of travel at highway segments and calculated using the following equation. Table 4.11 and 4.12 displays the first eighteen hazardous locations on Yangon-Mandalay expressway performed by Accident Rate Method.

$$R=\frac{(A\times 1{,}000{,}000)}{(365\times T\times V\times L)} \tag{4.5}$$

Where: R = accident rate for each segment (in accidents per million vehicle miles),

A = expected number of accidents for given analysis period,

T = time of analysis period (in years or fraction of years),

V = average annual daily traffic (AADT) during study period, and

L = length of highway segment (in miles).

Table 4.11 Output of *Accident Rate analysis* based on one mile interval for the most dangerous eighteen segments (*Yangon-Mandalay direction*)

Rank No	Beginning Mile	Ending Mile	Accident Rate
1	125	126	2.39
2	275	276	1.10
3	311	312	0.96
4	113	114	0.95
5	242	243	0.95
6	276	277	0.90
7	352	353	0.90
8	84	85	0.89
9	307	308	0.86
10	357	358	0.80
11	281	282	0.79
12	330	331	0.76
13	316	317	0.76
14	39	40	0.74
15	101	102	0.71
16	255	256	0.71
17	185	186	0.70
18	256	257	0.70

Table 4.12 Output of *Accident Rate analysis* based on one mile interval for the most dangerous eighteen segments (*Mandalay -Yangon direction*)

Rank No	Beginning Mile	Ending Mile	Accident Rate
1	115	116	0.96
2	352	353	0.94
3	117	118	0.83
4	83	84	0.77
5	242	243	0.69
6	116	117	0.66
7	89	90	0.64
8	246	247	0.64
9	129	130	0.63
10	184	185	0.63
11	64	65	0.63
12	223	224	0.63
13	294	295	0.62
14	218	219	0.62
15	125	126	0.62
16	67	68	0.60

Table 4.12 Output of *Accident Rate analysis* based on one mile interval for the most dangerous eighteen segments (*Mandalay -Yangon direction*) (Cont.)

Rank No	Beginning Mile	Ending Mile	Accident Rate
17	301	302	0.59
18	27	28	0.59

4.7.3 Rate Quality Control Method

In this method, the average accident rate (Ra) was determined by using the accident rate computed in the Accident Rate Method. The average accident rate (Ra) was calculated dividing the sum of accident rate of every segment by the total number of segments. The critical accident rate (Rc) for highway segment was then computed by the following equation.

$$Rc = Ra + K\left[\frac{Ra}{E}\right]^{0.5} + \frac{1}{2E} \qquad (4.6)$$

Where: Rc = critical accident rate for highway segment (accidents per million vehicle-miles),

Ra = average accident rate for all highway segments of similar characteristics,

E = (365*T*V*L)/1,000,000, million vehicle-miles of travel on the highway segment during the study period and

K =a probability factor determined by the desired level of significance for the equation. K is representing 95% confidence level and its value is 1.645.

The computed critical accident rate (Rc) was compared to the actual accident rate (R). If the actual accident rate higher than the critical rate, the road segment can be considered to be improved. The highway segments were ranked based on their (R/Rc) ratio in descending order. The segment with highest (R/Rc) value was determined as more dangerous and ranked as first hazardous location. Table 4.13 and 4.14 represents the first eighteen hazardous locations on Yangon-Mandalay expressway performed by Rate Quality Control Method.

Table 4.13 Output of *Rate Quality Control analysis* based on one mile interval for the most dangerous eighteen segments (*Yangon-Mandalay direction*)

Rank No	Beginning Mile	Ending Mile	R/Rc
1	125	126	2.2
2	275	276	1.1
3	276	277	1.0
4	311	312	1.0
5	330	331	1.0
6	316	317	1.0
7	84	85	0.9
8	242	243	0.9
9	255	256	0.9
10	256	257	0.9
11	357	358	0.9
12	352	353	0.8
13	307	308	0.8
14	185	186	0.8
15	113	114	0.8
16	39	40	0.8
17	281	282	0.7
18	101	102	0.7

Table 4.14 Output of *Rate Quality Control analysis* based on one mile interval for the most dangerous eighteen segments (*Mandalay –Yangon direction*)

Rank No	Beginning Mile	Ending Mile	R/Rc
1	115	116	0.8
2	83	84	0.7
3	117	118	0.7
4	352	353	0.7
5	89	90	0.6
6	64	65	0.6
7	116	117	0.6
8	67	68	0.6
9	129	130	0.6
10	27	28	0.6

Table 4.14 Output of *Rate Quality Control analysis* based on one mile interval for the most dangerous eighteen segments (*Mandalay –Yangon direction*) (Cont.)

Rank No	Beginning Mile	Ending Mile	R/Rc
11	184	185	0.6
12	242	243	0.5
13	125	126	0.5
14	188	189	0.5
15	60	61	0.5
16	246	247	0.5
17	8	9	0.5
18	153	154	0.5

4.7.4 Combined Method

In Combined Method, highway locations were ranked based on their hazardous index. Hazardous Index (HI) was calculated by taking sum all ranked numbers of previous three methods (Accident Frequency Method, Accident Rate Method and Rate Quality Control Method). After calculating hazardous indexes on each of road segment, all road segments were ranked in ascending order; meaning that the segment with lowest value of (HI) was determined as more dangerous location and ranked first. Table 4.15 and 4.16 describes the most dangerous eighteen hazardous locations on Yangon-Mandalay expressway performed by Combination Method. Hazardous Index (HI) was accomplished by the following equation.

$$HI = (F_{rank}+R_{rank}+Q_{rank}) / 3 \tag{4.7}$$

Where: HI = hazardous index,

F_{rank} = rank of location by accident frequency method,

R_{rank} = rank of location by accident rate method and

Q_{rank} = rank of location by rate quality control method.

The denominator value of "3" represents the number of methods whose ranks are summed in the numerator. Combined method is ranking of high accident concentration of each road segments by using equal weight into the individual method. This method enhances more accuracy than using other method alone. In this regards, a new approach, using different weight into the individual method, might be the one which can boost the level of accuracy for the future study of identifying hazardous locations.

Table 4.15 Output of *Combined Method* based on one mile interval for the most dangerous eighteen segments (*Yangon-Mandalay direction*)

No	Beginning Mile	Ending Mile	F_{rank}	R_{rank}	Q_{rank}	HI
1	125	126	1	1	1	3
2	275	276	4	2	2	8
3	84	85	3	8	7	18
4	113	114	2	4	15	21
5	276	277	13	6	3	22
6	242	243	11	5	8	24
7	311	312	21	3	4	28
8	358	359	10	10	11	31
9	39	40	5	14	16	35
10	101	102	6	15	18	39
11	185	186	15	17	14	46
12	353	354	29	7	12	48
13	281	282	23	11	17	51
14	307	308	33	9	13	55
15	63	64	7	23	28	58
16	75	76	8	24	29	61
17	255	256	38	16	9	63
18	105	106	9	26	30	65

Table 4.16 Output of *Combined Method* based on one mile interval for the most dangerous eighteen segments (*Mandalay*-(*Yangon direction*)

No	Beginning Mile	Ending Mile	F_{rank}	R_{rank}	Q_{rank}	HI
1	115	116	1	1	1	3
2	83	84	2	4	2	8
3	117	118	3	3	3	9
4	89	90	4	7	5	16
5	353	354	13	2	4	19
6	116	117	8	6	7	21
7	64	65	5	11	6	22
8	129	130	10	9	9	28
9	67	68	6	16	8	30
10	184	185	10	11	11	32
11	27	28	7	18	10	35
12	242	243	23	5	12	40
13	125	126	14	15	13	42
14	60	61	9	28	15	52
15	188	189	19	19	14	52
16	8	9	12	29	17	58
17	153	154	20	21	18	59
18	246	247	35	8	16	59

In summary, crash frequencies contribute as a major factor in Accident Frequency Method to identify hazardous locations. In addition to crash frequencies, traffic volumes also take effectiveness of Accident Rate Method and Rate Quality Control Method. Therefore, the ranking result of these two methods is nearly similar. However, their results are different from the result obtained from Accident Frequency Method. Lastly, the combined method (taking weighted average of ranking between Accident Frequency Method, Accident Rate Method and Rate Quality Control Method) provides an effective ranking result which is rather compatible to the results of each individual method.

CHAPTER 5

CONCLUSION

The study focused on two approaches of finding; examining relationships between traffic crash and road characteristics and identifying hazardous locations based on expected numbers of crash on Yangon-Mandalay Expressway in Myanmar. A crash prediction model was developed to determine relationships between traffic crash and road characteristics. On this subject, Negative Binomial regression, a family of Generalized Linear Model, was developed to estimate the model parameters. The sign and value of model parameters displayed the relationships between crash frequency and road characteristics on Yangon-Mandalay expressway. It was found out that five predictor variables including presence of bridge, Average Daily Traffic, combination of horizontal curve and slope, percent downgrade and presence of village settlement have a strong relationship with traffic crash happening.

In identifying hazardous location; crash frequencies predicted by the prediction model were refined by the Empirical Bayes (EB) estimation method. The use of EB approach enhanced precision for the estimation of crash frequency. After determination the crash frequencies on each road segment, three methods (Accident Frequency Method, Accident Rate Method and Rate Quality Control Method) were individually performed. In these methods, road segments were ranked based on their level of risk; meaning that the segment with highest crash frequency was ranked first. Finally these three methods were combined by taking sum all ranks number of each road segment. And the result was defined as danger index. The segment with smallest value of danger index was denoted more critical and ranked as first hazardous location.

In order to identify more reliable and true hazardous locations, all possible risk factors that contribute to crash occurrences should be considered in developing crash prediction model. Generally, traffic crashes take place due to human related factor, road related factor, vehicle related factor, environmental related factor and their interrelationships. Currently, traffic crash datasets including all kind of risk factors are not available in Myanmar. Therefore, only some road characteristics variables were utilized in the study for developing crash prediction model. All road segments on Yangon-Mandalay expressway were ranked depending on their degree of risk to identify hazardous locations.

The finding can be used in prioritizing road safety improvement program under limited budget condition. Furthermore, it will contribute to understand relationships between road characteristics and traffic crashes and to provide proper treatments in all hazardous locations on Yangon-Mandalay Expressway.

Recommendations

With regard to the finding of examining relationships between crash frequency and road characteristics on Yangon-Mandalay expressway, the following recommendations can be provided.

A combination of education, enforcement and engineering measures will be required to eliminate the crash frequencies on Yangon-Mandalay Expressway. The geometric improvements for traffic safety on road sections where horizontal alignment combined with vertical alignment include: (1) widening the lane and shoulder width on sharp horizontal curves, (2) increasing the amount of super elevation (up to maximum allowable rate), (3) increasing the distance of road side clear zone, (4) increasing the value of side friction (up to maximum allowable value) on every downgrade curve site, (5) delineating the pavement at all hazard locations to provide visual information to road users, and (6) providing warning by the use of traffic control devices to reduce vehicle speed limit at every approach to the sharp curve and downgrade locations.

Safety improvements on the existing bridges on Yangon-Mandalay Expressway includes: (1) providing speed reduction signs at every approach to the skew bridge and the bridge consists of complex geometric approach, (2) providing bridge approach guard rails with proper transition and end treatment, and (3) black and yellow hazard marking should be supplied at the areas surrounding the bridge to catch the driver's attentions on the presence of it along the expressway.

The influence of pedestrians and motorbikes on Yangon-Mandalay Expressway can be determined as the lack of enforcing road rules and road safety education to users. To overcome this problem, road authorities fundamentally need to implement how to educate and enforce road safety to the local people who live near the expressway. In addition, there is a need to be constructed other minor roads; i.e., frontage road, roads connect village to village and village to town with interchange if it is required, so that road users are needless to assess the expressway and cross directly any other traffic streams.

Research Publication

Parts of this work were published and presented in the following conference and journal;

Mon, C. T., Pueboobpaphan, R., & Ratanavaraha, V. (2016). *Identifying Hazardous Locations based on Expected Crash Frequency on Yangon-Mandalay Expressway in Myanmar*.

Paper presented at the Proceeding of 9th ATRANS Symposium: Young Researcher Forum 2016 "Transportation for a Better Life: Safe and Smart Cities", Bangkok, Thailand.

Mon, C. T., Pueboobpaphan, R., & Ratanavaraha, V. (2016). *Examining Relationship between Accident Occurrences and Road Characteristics on Yangon - Mandalay Expressway in Myanmar.*

International Journal of Building, Urban, Interior and Landscape Technology (BUILT), Vol (7).

REFERENCES

AASHTO, (2010) **Highway Safety Manual**. 1 ed. Vol. 1: American Association of State Highway and Transportation Officials.

Ackaah, W. and M. Salifu, (2011). Crash prediction model for two-lane rural highwaysin the Ashanti region of Ghana. **IATSS Research**, Vol.35, No.1: p. pp.34-40.

Ahsan, H. M., Newaz, M. S., Alam, A. K. M. S., & Alam, M. (2011). Application of GIS in Hazardous Road Location (HRL) Identification. **4th Annual Paper Meet and 1st Civil Engineering Congress. 2011.**

Amgalan, E., Jigjjav, M., Tsevegmid, T., Ichinnorov, B., & Purevdorj, C. (2013). Evaluation of traffic safety conditions where accidents are frequently occurred (road section between Ulaanbaatar-Baganuur). **Proceedings of the Eastern Asia Society for Transportation Studies, 9.**

Andersen, C. S., Olesen, A. V., & Bolet, L. (2013). Influenceof Road Characteristics on Density of Accidents on Secondary Rural Roads in Denmark. **from International Co-operation on Theories and Concepts in Traffic Safety** (ICTCT).

Arm, A. (2010). Effective safety factors on horizontal curves of two-lane highways. **Journal of Applied Science**,10,2814-2822.

Bauer, K. M., & Harwood, D. W. (2013). Safety Effects of Horizontal Curve and Grade Combinations on Rural Two-Lane Highways. **TRB 2013 Annual Meeting.**

Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. **Accident Analysis and Prevention**, 39(4), 657-670. doi:10.1016/j.aap.2006.10.012.

Ceunynck, T.D., et al., (2011). Explanatory models for crashes at high-risk locations**. Proceedings of the 24th ICTCT Workshop**, pp.1-17.

Dissanayake, S. and I. Ratnayake, (2006). Statistical modelling of crash frequency on rural freeways and two-lane highways using negative binomial distribution. **Advances inTransportation Studies an international Journal Section B 9.**

Eenink, R., (SWOV), M. R., (TOI), R. E., Cardoso, J., (LNEC), S. W., & (KfV), C. S. (2008). Accident Prediction Models and Road Safety Impact Assessment. **Sustainable Surface Transport** (pp. 20).

Elvik, R., (2008). A survey of operational definitions of hazardous road locations in some European countries. **Accident Analysis and Prevention**, Vol.40,No.6: p. pp.1830-5.

Gharaybeh,F.A., (1991). Identification of accident-prone locations in Greater Amman. **Transportation Research Board 1318, TRB, National Research Council, Washington D.C**., PP.70-74.

Greibe, P. (2003). Accident prediction models for urban roads. **Accident Analysis and Prevention**, 35, 273–285

Hanno, D. (2004). Effect of the combination of horizontal and vertical alignment on road safety. **(Master of Applied Science), University of British Columbia**.

Hauer, E., et al., (2002). Estimating Safety by the Empirical Bayes Method: A Tutorial. **Transportation Research Record No. 1784**, pp.126-131.

Kibar, F. T., Celik, F., & Aytac, B. P. (2013). An accident prediction model for divided highways: **a case study of Trabzon coastal divided highway**. 1, 711-719. doi: 10.2495/ut130571.

Kordani, A. A., & Molan, A. M. (2014). The effect of combined horizontal curve and longitudinal grade on side friction factors. K**SCE Journal of Civil Engineering,** *19*(1), 303-310. doi:10.1007/s12205-013-0453-3.

Kumar, C. N., Parida, M., & Jain, S. S. (2013). Poisson family regression techniques for prediction of crash counts using Bayesian Inference. **Procedia -Social and Behavioral Sciences**, 104, 982-991. doi:10.1016/j.sbspro.2013.11.193.

Lee, S., & Lee, Y. (2013). Calculation Method for Sliding-window Length: A Traffic Accident Frequency Case Study. **Proceedings of the Eastern Asia Society for Transportation Studies, 9.**

Lord, D., Geedipally, S. R., & Guikema, S. D. (2010). Extension of the application of Conway-Maxwell-Poisson models**:** Analyzing traffic crash data exhibiting under-dispersion. **Risk Analysis**, 30, 1268-1276. doi:10.1111/j.1539-6924.2010.01417.

Miaou, S.-P., & Lum, H. (1993). Modeling vehicle accidents and highway geometric design relationships. **Accident Analysis & Prevention**, 25(6), 689-709. doi:10.1016/0001-4575(93)90034-t.

Mohammed, A., S. Y. Umar, D. S., & Ahmad, T. Y. (2015). The Effect of Pavement Condition on Traffic Safety: A Case Study of Some Federal Roads in Bauchi

State. **IOSR Journal of Mechanical and Civil Engineering**, 12(3 Ver. I (May. - Jun. 2015). doi: 10.9790/1684-1231139146.

Mohammadi, M. A., Samaranayake, V. A., & Bham, G. H. (2014). Crash frequency modeling using negative binomial models: An application of generalized estimating equation to longitudinal data. **Analytic Methods in Accident Research,** *2*, 52-69. doi: 10.1016/j.amar.2014.07.001.

Naznin, F., Currie, G., Logan, D., & Sarvi, M. (2016). Application of a random effects negative binomial model to examine tram-involved crash frequency on route sections in Melbourne, Australia. **Accident Analysis & Prevention**, 92, 15-21. doi: 10.1016/j.aap.2016.03.012.

Psarianos, B., Kontaratos, M., & Giotis, A. (1994). Minimum Horizontal Curve Radius as a Function of Grade Incured By Vehicle Motion In the Driving Mode. **Transportation Research Record, Journal of the Transportation Research Board.**

Rahman, M., & Newaz, S. (2013). Development of a GIS based hazardous road location identification system for National Highways of Bangladesh. **TOJSAT : The Online Journal of Science and Technology,** 3(2), 72-85.

Ratanavaraha, V. and C. Amprayn, (2003). Causative highway accident factors of the expressway system in Thailand. **Journal of the Eastern Asia Society for Transportation Studies,** Vol.5.

Saffarzadeh, M., & Pooryari, M. (2005). Accident Prediction Model Based on Traffic and Geometric Design Characteristics. **International Journal of Civil Engineerng, 3, 112-119.**

Shin, K., (2008). Empirical bayes method in the study of traffic safety accounting for spatial and temporal heterogeneity , **Arizona State University: United States**.

Smith, B.L., and Lamm,R.(1994). Coordination of horizontal and vertical alignment withregard to highway esthetics,**Transportation Research Record: Journal of the Transportation Research Board**,1994,1445, pp. 73-85.

Utainarumol, S. and J. Robert E. Stammer, (1999). An evaluation of methods for identifying hazardous highway locations, **Journal of the Eastern Asia Society for Transportation Studies**, Vol.3, No.1.

Valentová, V., J. Ambros, and Z. Janoška, (2014). A comparative analysis of identification of hazardous locations in regional rural road network. **Advances in Transportation Studies an international Journal, Section B 34, 2014.**

World Health Organization, H. (2013). **Globl Status Report**.

Y. Hassan, M. A., & S. M. Easa, M. A. (2003). Effect of Vertical Alignment on Driver Perception on Horizontal Curve. **Journal of Transportation Engineering,**129(4),399-407.doi:10.1061//ASCE/0733-947X/2003/129:4/399.

Zegeer, C.V., (1986). Synthesis of Highway Practice 128: Methods for identifying hazardous highway elements, **Transportation Research Board of the National Academics, Washington, *D.C.***

Zou, Y., Wu, L., & Lord, D. (2015). Modeling over-dispersed crash data with a long tail: Examining the accuracy of the dispersion parameter in Negative Binomial models. **Analytic Methods in Accident Research**,*5-6*,1-16. doi:10.1016/j.amar.2014.12.002.

BIOGRAPHY

Cho Thet Mon was born on February 1, 1981 in Myaungmya Township, Ayeyawaddy Region, Myanmar. She received her Bachelor’s degree with honors in Civil Engineering from Hmawbi Technological University in Myanmar. After graduation, she worked at Ministry of Construction as a Junior Engineer. Then she was awarded a full-time scholarship from TICA (Thailand International Development Cooperation Agency) to continue a Master’s degree in Transportation Engineering at Suranaree University of Technology, Thailand.